Comparing GARCH Volatility Distributions
To Implied Volatility Distributions

ECO2408

Introduction

The volatility of asset returns is of crucial importance in financial markets. Volatility serves as a measure of risk faced by economic agents who purchase an asset. Consequently, volatility is an integral part of many asset pricing models. Specifically, volatility is fundamental when using the Black-Scholes option pricing model. In fact, the only parameter that is unobservable in the Black-Scholes model is volatility. Therefore, the estimation of volatility is critical in order to use the model in practice. One of the most popular ways of modeling volatility is a generalized autoregressive conditional heteroscedastic (GARCH) model. The purpose of this study is to examine how well a volatility defined by a GARCH process matches the implied volatility observed in the market. The premise is that if we can match the GARCH volatility to the implied volatility we can justify the use a GARCH model to estimate future volatilities and consequently option prices. The paper will start out explaining the main underlying concepts of the Black-Scholes model and the GARCH model. Then a GARCH model will be formulated using maximum likelihood estimation based on Microsoft equity returns. The volatility estimates from the GARCH model will be compared to the implied volatility of Microsoft stock using nonparametric procedures.

The Black-Scholes-Merton Option Pricing Model

To understand the Black-Scholes-Merton pricing model it is helpful to first define a model of stock price behaviour. Let us assume that agents require an expected rate of return on their investment (i.e. a percentage of the underlying stock price) and that this return is constant over time. If S is the stock price at time t, the expected drift rate in S should be assumed to be μ S, where μ is a constant parameter representing the expected rate of return on the stock. Thus, the expected change in over some interval δ t is μ S. In addition, if we assume that volatility is constant over time we can arrive at the following model:

$$\delta S = \mu S \delta t + \sigma S \delta z \text{ where z is a Wiener process} \qquad (1)$$

From the properties of a Wiener process we know that z has the following properties:

$\delta z = \varepsilon\sqrt{\delta t}$ where ε is a random drawing from a standard normal distribution, i.i.d. N(0,1).

$E(\delta z) = 0$

$Var(\delta z) = E(\delta z)^2 = \delta t$

Moving from discrete time to continuous time our model is now:

$$dS = \mu S dt + \sigma S dz \Rightarrow \frac{dS}{S} = \mu dt + \sigma dz \Rightarrow \frac{dS}{S} = \mu dt + \sigma\varepsilon\sqrt{dt} \qquad (2)$$

The model of stock price behaviour developed above is known as a geometric Brownian motion. It is one of the most popular stochastic processes assumed for a non-dividend paying stock. The left hand side of equation (2) is the return of the stock over an infinitesimal time interval. The term μdt is the expected value of this return and the term $\sigma\varepsilon\sqrt{dt}$ is the stochastic component of the return. The variance of the stochastic component (and therefore the whole return) is $\sigma^2 dt$. From these results we can describe the distribution of stock returns as:

$$\frac{dS}{S} \sim N(\mu dt, \sigma^2 dt) \qquad (3)$$

Another important concept in deriving the Black-Scholes-Merton model is Itô's lemma, developed by Itô (1951). A rigorous proof of this result is left out of this paper, for more detail consult the references. Suppose that the value of a variable x follows the Ito process,

$dx = a(x,t)dt + b(x,t)dz$ where dz is a Wiener process and a and b are functions of x and t

The variable x has a drift rate of a and a variance of b^2. Itô's lemma shows that a function G of x and t follows the process,

$$dG = \left(\frac{\partial G}{\partial x}a + \frac{\partial G}{\partial t} + \frac{1}{2}\frac{\partial^2 G}{\partial x^2}b^2\right)dt + \frac{\partial G}{\partial x}bdz$$, where dz is the same Wiener process as above (4)

The Black-Scholes-Merton differential equation is developed using a geometric Brownian motion as described above in (2),

$$dS = \mu S dt + \sigma S dz \qquad (5)$$

Using Itô's lemma and defining $G = \ln S$,

$$\ln\left(\frac{S_T}{S_0}\right) \sim N\left[\left(\mu - \frac{\sigma^2}{2}\right)T, \sigma\sqrt{T}\right] \Rightarrow \ln S_T \sim N\left[\ln S_0 + \left(\mu - \frac{\sigma^2}{2}\right)T, \sigma\sqrt{T}\right] \qquad (6)$$

where S_T is the stock price at some future price T, and S_0 is the stock price at time zero. From (6), S_T is lognormally distributed since ln S_T is normally distributed.

Suppose that f is the price of a derivative contingent on S. The variable f must be some function of S and t. Utilizing Itô's lemma,

$$df = \left(\frac{\partial f}{\partial S}\mu S + \frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial S^2}\sigma^2 S^2\right)dt + \frac{\partial f}{\partial S}\sigma S dz \qquad (7)$$

Recall that the Wiener process (dz) underlying f and S are the same. Thus, by choosing an appropriate position in the stock and the derivative, the Weiner process can be eliminated. Specifically we can short the derivative and take a long position in the stock. The value of this position can is therefore,

$$\Pi = -f + \frac{\partial f}{\partial S}S \Rightarrow d\Pi = -df + \frac{\partial f}{\partial S}dS \qquad (8)$$

Substituting in (5) and (7),

$$d\Pi = \left(-\frac{\partial f}{\partial t} - \frac{1}{2}\frac{\partial^2 f}{\partial S^2}\sigma^2 S^2\right)dt$$, where Π is the value of the position (9)

Notice that dz does not appear in (9). Consequently, the position must be riskless over some very small period of time. Using a arbitrage argument the following relation can be asserted,

$$d\Pi = r\Pi dt \text{, where r is the risk-free rate} \qquad (10)$$

If the above relationship is not true then arbitrageurs could make a riskless profit by investing in either the risk-free asset or the portfolio, which ever provides a greater return. Substitution (8) and (9) into (10),

$$\left(\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial S^2}\sigma^2 S^2\right)dt = r\left(f - \frac{\partial f}{\partial S}S\right)dt \Rightarrow \frac{\partial f}{\partial t} + rS\frac{\partial f}{\partial S} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 f}{\partial S^2} = rf \qquad (10)$$

Equation (10) is the Black-Scholes-Merton differential equation. Notice that (10) is independent of μ. This means that the Black-Scholes-Merton differential equation is independent of risk preferences, thus, any set of risk preferences can be used to evaluate f. To assume that preferences are risk neutral greatly simplifies the analysis of derivatives.

The solution to equation (10) depends on the type of derivative that is being valued. For example, to value a European call option (c) the key boundary condition to (10) is,

$$f = \max(S - K, 0) \text{ when t=T, where K is the strike price}$$

Obviously, the boundary condition will change based on the type of option being valued. For example, the boundary condition for a European put (p) option is,

$$f = \max(K - S, 0) \text{ when t=T}$$

Based on the two boundary conditions specified above the solution to (10) is,

$$c = S_0 N(d_1) - Ke^{-rT}N(d_2) \qquad p = Ke^{-rT}N(-d_2) - S_0 N(-d_1) \qquad (11)$$

where

$$d_1 = \frac{\ln(S_0/K) + (r - \sigma^2/2)T}{\sigma\sqrt{T}} \qquad d_2 = \frac{\ln(S_0/K) + (r - \sigma^2/2)T}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T}$$

For a rigorous proof of the above solution please consult the references, Black (1973). In equations (11), T is the time to maturity measured in years, r is a continuous compounded risk-free return per year, and σ is the volatility of the stock measured in years.

Notice that in the derivation of (11) volatility is taken as a constant. That is it does not change with time. In addition, notice that volatility is the only parameter that can not be directly observed from the market. It is therefore necessary to estimate volatility empirically in order to use the Black-Scholes model in practice. A popular way of modeling volatility is using an ARCH processes so that volatility is allowed to vary over time. A we will see later, the implied volatility fluctuates over time, thus, it is important to account for this feature when estimating volatility.

ARCH Specification

Let ψ_t represent the information set at time t

$$Y_t = E[Y_t | \Psi_{t-1}] + \varepsilon_t$$

$$E(\varepsilon_t | \Psi_{t-1}) \sim N(0, \sigma^2) \tag{12}$$

Although (12) implies that the unconditional variance of ε_t is the constant σ^2, the conditional variance of ε_t could change over time. Letting the square of ε_t follow an AR(m) process:

$$\varepsilon_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \upsilon_t \text{, where } \upsilon_t \text{ is a white noise process: i.i.d. } E(\upsilon_t) = 0 \quad E(\upsilon_t^2) = \lambda^2 \tag{13}$$

An alternative representation for the above process that imposes slightly stronger assumptions about the serial dependence of ε_t,

$$\varepsilon_t = h_t z_t \text{, where } z_t \text{ is an i.i.d. sequence with } E(z_t) = 0 \text{ and } E(z_t^2) = 1 \tag{14}$$

$$h_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 \tag{15}$$

Substituting (14) and (15) into (13),

$$\upsilon_t = h_t^2 (z_t^2 - 1) \tag{16}$$

Note that the unconditional variance of υ_t ($E(\upsilon_t^2) = \lambda^2$) is assumed to be constant, however, the conditional variance will change over time.

A white noise error process ε_t satisfying (13) is described as an autoregressive conditional heteroscedastic process of order p, denoted ε_t ~ARCH(p). This class of process was introduced by Engle (1982). In order for h_t^2 to be covariance –stationary,

$$1 - \alpha_1 z - \alpha_2 z^2 - \ldots - \alpha_p z^p = 0 \tag{17}$$

the roots of (17) must lie outside the unit circle. If all α_i are non-negative then we can say this is equivalent to,

$$\alpha_1 + \alpha_2 + \ldots + \alpha_p < 1 \tag{18}$$

Satisfying these conditions result in the unconditional variance of ε_t is given by,

$$E(h_t^2) = \frac{\omega}{1 - \alpha_1 - \alpha_2 \ldots - \alpha_p} \tag{19}$$

GARCH Specification

The generalized ARCH (GARCH) model of Bollerslev (1986) is based on an infinite ARCH specification and reduces the number of estimated parameters. Letting the conditional variance depend on an infinite number of lags of ε_{t-1}^2,

$$h_t^2 = \omega + \sum_{i=1}^{\infty} \pi_i L^i \varepsilon_t^2 \tag{20}$$

Parameterize $\sum_{i=1}^{\infty} \pi_i L^i$ as the ratio of two finite-order polynomials,

$$\sum_{i=1}^{\infty} \pi_i L^i = \frac{\alpha(L)}{1-\beta(L)} = \frac{\alpha_1 L^1 + \alpha_2 L^2 + \ldots + \alpha_p L^p}{1 - \beta_1 L^1 - \beta_2 L^2 - \ldots - \beta_q L^q} \tag{21}$$

Multiplying equation (21) by $1-\beta(L)$,

$$[1-\beta(L)]h_t^2 = [1-\beta(L)]\omega + \alpha(L)\sum_{i=1}^{\infty} \pi_i L^i \varepsilon_t^2$$

$$h_t^2 = \delta + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{q} \beta_i h_{t-i}^2 \text{ , where } \delta = [1-\beta_1-\beta_2-\ldots-\beta_q]\omega \tag{22}$$

Expression (22) is the generalized autoregressive conditional heteroscedasticity model, denoted $\varepsilon_t \sim GARCH(p,q)$ proposed by Bollerslev (1986). To satisfy that h_t^2 is covariance-stationary if,

$$\sum_{i=1}^{m} \alpha_i + \sum_{i=1}^{m} \beta_i < 1 \text{, where } m \equiv \max(p,q) \tag{23}$$

Assuming this condition holds, the unconditional mean of h_t^2 is,

$$E(h_t^2) = \frac{\delta}{1 - \sum_{i=1}^{m} \alpha_i - \sum_{i=1}^{m} \beta_i} \tag{24}$$

To ensure that the conditional variance h_t^2 is nonnegative $\delta \geq 0$. Also requiring $\alpha_i \geq 0$ and $\beta_i \geq 0$ may seem sensible to ensure that the conditional variance is nonnegative. However, Nelson and Cao (1992) show that these conditions are sufficient but not necessary to ensure the nonnegativity of h_t^2.

Maximum Likelihood Estimation

ARCH and GARCH models are most often estimated by maximizing the likelihood function (ML). The logic of ML is to interpret the density as a function of the parameter set, conditional on a set of sample outcomes. Assuming the ε_t's are i.i.d. and normally distributed with zero mean and variance h_t^2, the log-likelihood function for a sample of T observations is,

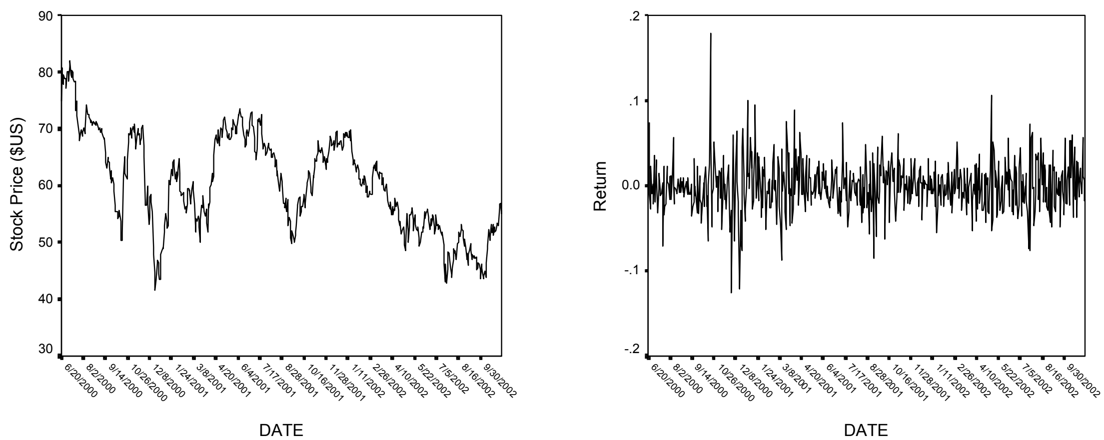
$$L(\theta) = -\frac{T}{2}\log(2\pi) - \frac{T}{2}\log(h_t^2) - \frac{1}{2}\sum_{t=1}^{T} \varepsilon_t^2 (h_t^2)^{-1} \tag{25}$$

Where θ is the set of parameters to be estimated. To maximize this function we will set up the first order conditions and solve the resulting set of non-linear, recursive equations via some numerical optimization algorithm. The Berndt, Hall, Hall and Hausman (1974) algorithm is used in this case to maximize the log-likelihood function.

Data Sources & Description

Equity returns from Microsoft Corporation will be used in to estimate a GARCH process. The return will be calculated using daily prices of Microsoft stock, and implied volatilities will be calculated using call option prices over the period June 20, 2000 to November 8, 2002, representing 600 observations. Figure (1) shows the price and return of Microsoft stock over the period in question.

FIGURE 1



Stock prices were retrieved from NSDAQ through the use of the Telerate software in the Financial Lab at University of Toronto, Rotman School of Management. The Call option prices were also retrieved from CBOE though the use of Telerate Software. The 3-month LIBOR rates were retrieved from DataStream using the DataStream advance software.

The three month London Interbank Offer Rate (LIBOR) as the risk free rate. Banks and other large financial institutions tend to use LIBOR rather than the Treasury rate as the risk free rate. The reason is that financial institutions invest surplus funds in the LIBOR market and borrow to meet their short term funding requirements in this market. Consequently, LIBOR can be views as their opportunity cost of capital.

The 3 month LIBOR rate is quoted in quoted in yield, thus, it must be converted into a continuously compounded annual rate:

$$\text{LIBOR Yield}=\frac{360}{91}*(100-Y) \qquad \text{Y is the cash price}$$

The quarterly compound rate per annum (Rm) is calculate as follows:

$$Rm=\frac{(100-Y)}{Y}*\frac{365}{91}$$

The continuous compounding rate (Rc) is calculated as follows:

$$Rc=4*\ln(1+\frac{Rm}{4})$$

The RC in the model is defined as the risk free rate of return.

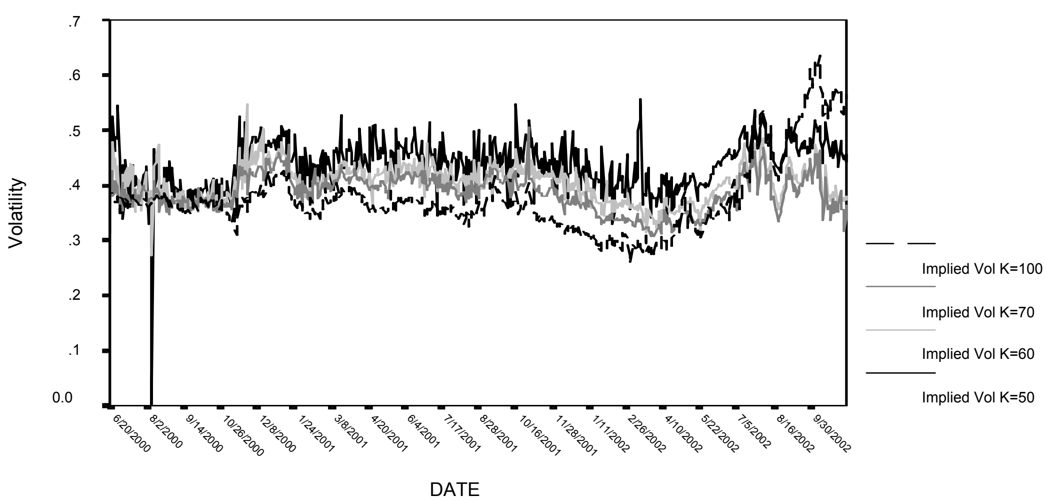
Implied Volatility:

The one parameter in the Black-Scholes pricing formulas that cannot be directly observed is the volatility of the stock price. The way to estimate the implied volatility is to keep the stock price, strike price, interest rate and time to maturity fixed, while adjusting the volatility until the Black-Scholes formula converges to the call price actually observed in the market.

Call Options:

Since short-term options have a small number of observations (they have a very short history), longer maturity call options must be used. A long-term equity anticipation security (LEAP) have relatively longer histories than regular options. LEAP (one and half year life span) call options that mature in 2003 with strike prices of \$50, \$60, \$70, and \$100 will be used to calculate implied volatility. It is possible that the distribution of implied volatility is dependent on strike price. Thus, a range of strike prices will be used to account for options that are in-the-money or out-of-the-money. The calculated implied volatilities are displayed in Figure (2).

FIGURE (2)



Analysis of Empirical Model

The unit root test for stationarity tries to discriminate series that are stationary from those that follow a non-stationarity (or integrated) process like a pure Random Walk. So the purposes of our GARCH model we had to test whether our variable r (i.e. equity returns) followed an integrated process of order zero or a random walk process. The reason why this is important is because we want make certain that our coefficients are consistent and that our statistical procedures are in fact the correct ones. The Augmented Dickey Fuller test is performed by estimated the unrestricted regression equations with lagged changes of endogenous variables so as to control for autocorrelated errors. The number of lags is contingent upon the frequency of the sample data. In our case, with daily observations, we used up to 30 lags of the endogenous variable. This procedure will give us a better estimate of the test statistic since we are allowing for the existence of serial correlation in the residuals. Once we have calculated the SSE of the unrestricted equation we then impose the restrictions of a random walk *without a drift* (row to one) and estimate the SSE of the restricted equation. Calculating a standard F test (using the restricted SSE and Unrestricted SSE) and comparing it to the D-F critical values will tell us if autocorrelation coefficient is significantly different from one. From table (1) we see that in out the autocorrelated coefficient in significantly different from one (i.e. less than one) at a 1%, 5% and 10% level of confidence. So we conclude that our variable r follows a stationary process. (Note: Since we are only using one variable, namely equity returns, we don't have to test for co-integration).

$$\Delta Y_t = \delta Y_{t-1} + \sum_{i}^{p} \alpha_i \Delta Y_{t-1} + u_t$$

H_0: $\delta = 0$ (nonstationary)

H_1: $\delta < 0$ (stationary)

TABLE 1

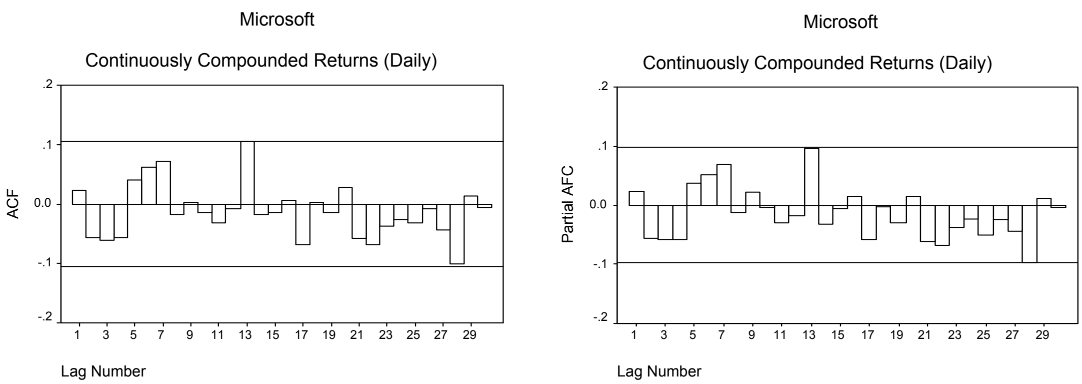
Augmented Dickey-Fuller test for unit root Number of obs = 568

	Test Statistic	1% Critical Value	5% Critical Value	10% Critical Value
		---------- Interpolated Dickey-Fuller ---------		
Z(t)	-5.618	-3.430	-2.860	-2.570

* MacKinnon approximate p-value for Z(t) = 0.0000

We can also see this result by looking at the autocorrelation functions and partial autocorrelation function, Figure 3. In the ACF as well as the PACF we see how each coefficient are not significantly different from zero. This means that equity returns (i.e. r) follows are white noise process and thus the price level (i.e. original model) follows a random walk. This can also be tested using the Piece-Box-Q test or DW test to check whether our coefficient is indeed close to zero.

FIGURE 3



EMPERICAL MODEL:

In the following we assume the conditional mean specification is,

$R_t = \kappa + \varepsilon_t$ where R_t represents daily continuously compounded returns (26)

$\varepsilon_t = z_t h_t$ with $E(z_t) = 0$ $Var(z_t) = 1$ and $\varepsilon_t \sim N(0, h_t^2)$ (27)

$$h_t^2 = \delta + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-1}^2 + \sum_{i=1}^{q} \beta_i h_{t-1}^2 \quad (28)$$

First, however, the ARCH test will be used to illustrate the validity of using an ARCH type model. Engle(1982) derived this test using the Lagrange multiplier principle.

$$\mathrm{var}(\varepsilon_t^2) = h_t^2 = \lambda_0 + \lambda_1 \varepsilon_{t-1}^2 + \lambda_2 \varepsilon_{t-2}^2 + \ldots + \lambda_p \varepsilon_{t-p}^2 \quad (29)$$

H_0: $\lambda_1 = \lambda_2 = \ldots = \lambda_p = 0$, in which case we have $\mathrm{var}(\varepsilon_t^2) = \lambda_0$, implying a homoscedastic error variance.

H_1: at least one $\lambda_1, \lambda_2, \ldots, \lambda_p \neq 0$

ARCH(1), ARCH(2), ARCH(3), ARCH(4), and ARCH(5) models are estimated for this test. That is p=5. The test statistic, nR^2, follows a χ^2 distribution with p degrees of freedom.

TABLE 2

	R^2	n	p	NR^2	p-value
ARCH(1)	0.0085	598	1	5.083	0.0241615
ARCH(2)	0.0126	597	2	7.5222	0.0232581
ARCH(3)	0.0237	596	3	14.1252	0.0027396
ARCH(4)	0.0244	595	4	14.518	0.0058128
ARCH(5)	0.0262	594	5	15.5628	0.0082097

For ARCH(1)-ARCH(5) there is sufficient evidence to reject H_0 at a 95% confidence level. Therefore, we can conclude that the error variance is serially correlated. This gives merit to using an ARCH type model.

The parameter estimates for the above model will be found using the maximum likelihood technique as described earlier in this paper. However, the order of GARCH(p,q) processes still needs to be addressed. To determine the order of the GARCH model, the Schwartz criterion has been shown to deliver consistent series models by Bollerslev, Chou and Kroner (1992). In addition, the Akaike and Hannan-Quinn criterion also produce similar results to the Schwartz criterion. The formulas for the above criterion along with their computations for a GARCH $p \in [1,3]$, and $q \in [1,3]$ are shown below. Minimizing the following criteria determines the values of p and q.

- Akaike $= -2\frac{\log L}{n} + 2\frac{k}{n}$
- Schwartz $= -2\frac{LogL}{n} + 2\frac{\log(k)}{n}$
- Hannan-Quinn $= -2\frac{\log L}{n} + 2\frac{k\log[\log(k)]}{n}$
- Shibata $= -2\frac{\log L}{n} + \log\left(\frac{n+2k}{n}\right)$

Where k is the number of parameters estimated, n is the sample size, and Log L is the log-likelihood value.

TABLE 3

	P	Q	N	Log Likelihood	Akaike	Schwartz	Hannan-Quinn
GARCH(1,1)	**1**	**1**	**599**	**-1267.363**	**4.244951586**	**4.23622469**	**4.235958387**
GARCH(2,1)	2	1	599	-1278.539	4.28560601	4.274285268	4.276856177
GARCH(3,1)	3	1	599	-1281.912	4.300207012	4.286156125	4.291857057
GARCH(1,2)	1	2	599	-1279.014	4.287191987	4.275871245	4.278442154
GARCH(2,2)	2	2	599	-1286.969	4.31709182	4.303040933	4.308741865
GARCH(3,2)	3	2	599	-1282.163	4.304383973	4.287508882	4.296571314
GARCH(1,3)	1	3	599	-1285.67	4.312754591	4.298703704	4.304404636
GARCH(2,3)	2	3	599	-1283.311	4.308217028	4.291341937	4.30040437
GARCH(3,3)	3	3	599	-1287.762	4.326417362	4.306649221	4.319261419

The results from Table (3) show that Schwartz, Akaike and Hannan-Quinn criterion all confirm that GARCH(1,1) is the optimal order of the model. In addition, Bollerslev, Chou and Kroner (1992) have found that the GARCH(1,1) model best captures the heteroscadasticity characteristics of financial data.

MODEL:

$R_t = \kappa + \varepsilon_t$ where R_t represents daily continuously compounded returns (30)

$\varepsilon_t = z_t h_t$ with $E(z_t) = 0$ $Var(z_t) = 1$ and $\varepsilon_t \sim N(0, h_t^2)$ (31)

$$h_t^2 = \delta + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}^2 \quad (32)$$

Maximum Likelihood Estimation

Sample: 2 to 600		Number of obs	=	599
		Wald chi2(.)	=	.
Log likelihood = -1267.363		Prob > chi2	=	.

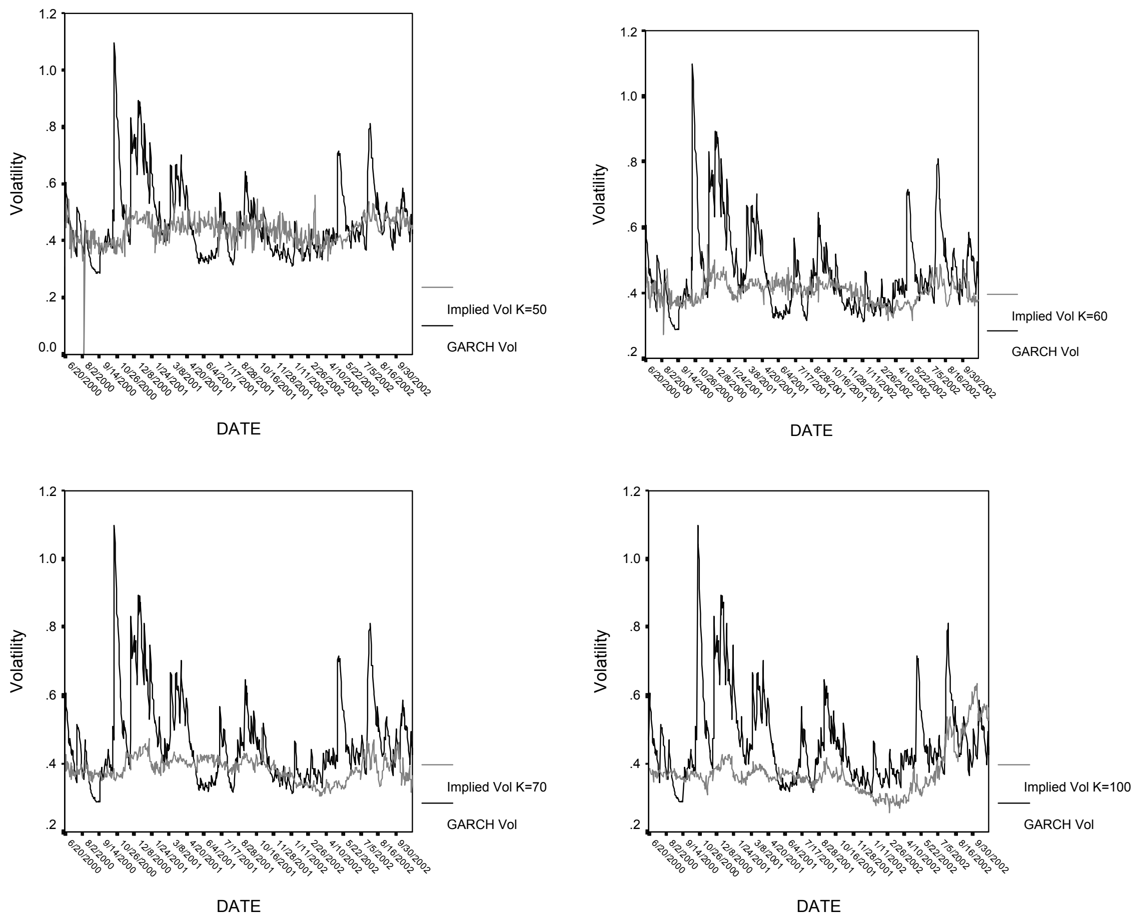
	Coef.	OPG Std. Err.	z	P>|z|	[95% Conf.	Interval]
κ	-.0003579	.0010635	-0.336	0.736	-.0024422	.0017265
GARCH						
α_1	.1246635	.0237754	5.243	0.000	.0780645	.1712624
β_1	.8282759	.034424	24.061	0.000	.7608062	.8957457
δ	.0000475	.0000175	2.708	0.007	.0000131	.0000819

The conditional variance, h_t^2, is measured in days. To compare the volatility produced by the GARCH model with the implied volatilities observed in the market we must convert h_t^2 so that it is measured in years (252 trading days per year). We will denote the yearly volatility produced by GARCH estimate as σ_{GARCH}, and the implied volatility will be denoted as $\sigma_{IMPLIED}$.

$$\sigma_{GARCH,t} = \sqrt{(h_t^2)(252)} \quad (33)$$

Figure (4) shows σ_{GARCH} compared to $\sigma_{IMPLIED}$ under strike prices of \$50, \$60,\$70 and \$100.

FIGURE 4



The above graphs seem to indicate that the GARCH volatility matches the implied volatility better for lower strike prices than for higher strike prices. To test this hypothesis formally we should test whether the distribution of σ_{GARCH} matches the distribution of $\sigma_{IMPLIED}$. In situations where the population distribution function (pdf) is know with certainty a goodness of fit test will suffice. However, in this situation the pdf of implied volatility is unknown, hence, an alternative testing procedure must be used. One way of dealing with this problem is to use test statistics whose pdf's remain the same, regardless of how the population sampled may change. Inference procedures having this sort of scope are called nonparametric. One of the simplest nonparametric tests is the sign test. It is appropriate in a paired-data situation where the normality if the distribution is in question. The basic approach is to test whether the number of data values that have increased is different from the number of data values that decreased. Under H_0 we have:

probability of an increase = the probability of a decrease

P(increase) = P(decrease)

Binomial distribution with $\pi = 0$

If $n \geq 25$

Let $d_i = x_1 - x_2 = \sigma_{GARCH,t} - \sigma_{IMPLIED,t}$

$n =$ number of $d_i \neq 0$

$T =$ number of positive d_i

Test statistic: $Z_{calc} = \dfrac{(T \pm 0.5) - 0.5n}{0.5\sqrt{n}}$, where T is adjusted for continuity

The basic draw back of the sign test is that it assumes the data set comes from a symmetric distribution. Since we do not have a clear idea of whether the distribution is indeed symmetric, an alternative procedure will be discussed. The Wilcoxon signed rank test is also appropriate for comparing distributions of paired-data where true functional form of the distribution is unknown. The Wilcoxon signed rank test is one of the most widely used nonparametric procedures, and dominates the sign test in terms of both size and power. Under H_0 we have the populations being identical.

If $n > 30$

Define $W = \min(W_+, W_-)$

where W_+ is the rank sum of the positive differences

where W_- is the rank sum of the negative differences

W is approximately normally distributed with $E(W) = \dfrac{n(n+1)}{4}$ and $\sigma_w = \sqrt{\dfrac{n(n+1)(2n+1)}{24}}$

$Z_{calc} = \dfrac{W - E(W)}{\sigma_w}$, with an upper critical value $-Z_{\alpha/2}$ and lower critical value $\dfrac{n(n+1)}{2}$

The results of both the Wilcoxon signed rank test, and the sign test are shown below in Table (4).

TABLE 4

Wilcoxon Signed Ranks Test

Ranks

		N	Mean Rank	Sum of Ranks
GARCH Volatility - Implied Volatility K=50	Negative Ranks	295	270.80	79885.00
	Positive Ranks	305	329.23	100415.00
	Ties	0		
	Total	600		
GARCH Volatility - Implied Volatility K=60	Negative Ranks	196	232.17	45506.00
	Positive Ranks	404	333.65	134794.00
	Ties	0		
	Total	600		
GARCH Volatility - Implied Volatility K=70	Negative Ranks	152	188.84	28703.00
	Positive Ranks	448	338.39	151597.00
	Ties	0		
	Total	600		
GARCH Volatility - Implied Volatility K=100	Negative Ranks	121	198.57	24027.00
	Positive Ranks	479	326.25	156273.00
	Ties	0		
	Total	600		

Test Statistics

	GARCH Volatility - Implied Volatility K=50	GARCH Volatility - Implied Volatility K=60	GARCH Volatility - Implied Volatility K=70	GARCH Volatility - Implied Volatility K=100
Z	-2.416	-10.510	-14.465	-15.566
Asymp. Sig. (2-tailed)	.016	.000	.000	.000

Sign Test

Frequencies

		N
GARCH Volatility - Implied Volatility K=50	Negative Differences	295
	Positive Differences	305
	Ties	0
	Total	600
GARCH Volatility - Implied Volatility K=60	Negative Differences	196
	Positive Differences	404
	Ties	0
	Total	600
GARCH Volatility - Implied Volatility K=70	Negative Differences	152
	Positive Differences	448
	Ties	0
	Total	600
GARCH Volatility - Implied Volatility K=100	Negative Differences	121
	Positive Differences	479
	Ties	0
	Total	600

Test Statistics

	GARCH Volatility - Implied Volatility K=50	GARCH Volatility - Implied Volatility K=60	GARCH Volatility - Implied Volatility K=70	GARCH Volatility - Implied Volatility K=100
Z	-.367	-8.451	-12.043	-14.574
Asymp. Sig. (2-tailed)	.713	.000	.000	.000

Sign Test

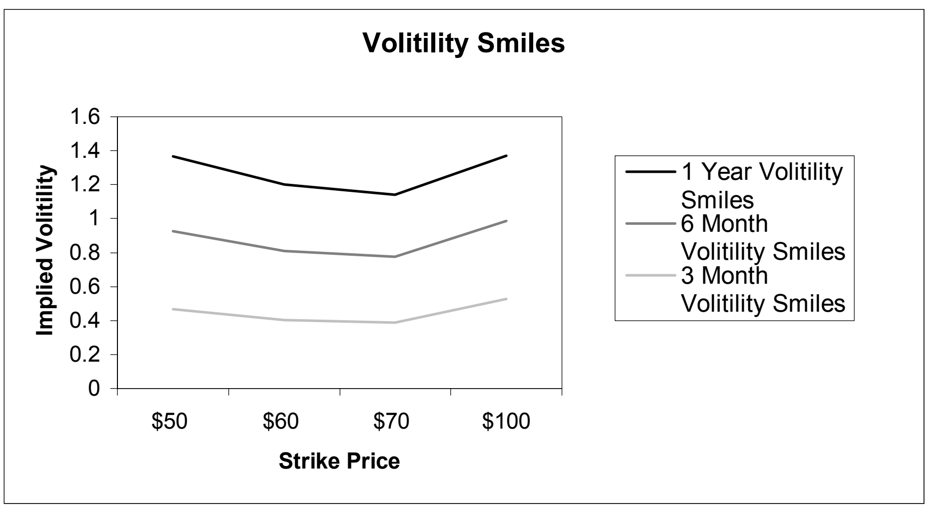
The results from Table (4) indicate that the distribution of GARCH volatility is different from the implied volatility under strike prices of \$50, \$60, \$70, and \$100. All test results, except for the sign test with a strike price of \$50, clearly reject H_0. Although, the statistical evidence suggests that the distributions are different under all the strike prices examined, there is an indication that the GARCH volatility distribution matches the implied volatility distribution better for lower strike prices. This result can be drawn from the fact that the Z_{CALC}'s decrease as strike price decreases for both the Wilcoxon signed rank test and the sign test.

Conclusion

A GARCH(1,1) model was used to estimate conditional variance on Microsoft stock over the period starting at June 20,2000 to November 8, 2002. The conditional variance was then compared to the implied volatility in the market to assess the validity of using a GARCH process to estimate volatility. Nonparametric tests were used to compare the distributions of the estimated GARCH volatility and the implied volatility in the market. The results indicate that the two distributions in question are not equal to each other. However, we can conclude that GARCH(1,1) conditional volatility matches the implied volatility distribution better for lower strike prices (this corresponds to deep out-of-the-money call options or deep in-the-money put options).

The GARCH model used in this study does not produce volatilities that are consistent with empirical implied volatilities. However, other ARCH-type models possess properties that may help to estimate volatility more accurately. This point can be illustrated with the aid of a volatility smile. A volatility smile is a plot of the implied volatility of an option as a function of its strike. Three volatility smiles are displayed in Figure (5) corresponding to time to maturity of less than 3 months, between 3 months and 6 months, and more than 1 year.

FIGURE 5



The volatility smile in Figure (5) indicates that the relationship between volatility, strike price and time to maturity is dynamic. If volatility were constant over time and strike price the volatility smiles should all be the same horizontal line. However, we can see that volatility becomes larger as the strike price progresses to extreme low or high values. Therefore, there is a greater probability of the stock price being extremely low or high. Furthermore, the curvature of the volatility smile seems to depend on the option maturity. The smile tends to be less pronounced as the option maturity increases. The above arguments suggest that the volatility of asset returns display asymmetric properties. In other words, positive innovations in asset prices are characteristic of greater volatility than negative innovations. The GARCH model prescribed in this study treats the volatility of innovations as symmetric, that is, positive and negative innovations have the same effect on volatility. Obviously, this model misspecification has a detrimental effect on the validity of the model developed in this study. In order to account for asymmetry, or the "leverage effect" as it has been termed, Nelson developed the exponential GARCH (EGARCH) model that accounts for both the size and the sign of lagged residuals in the conditional variance. Another popular GARCH extension is the fractionally integrated GARCH (FIGARCH) model developed by Baillie, Bollerslev, and Mikkelsen. Empirical studies often result in $\sum_{j=i}^{p}\alpha_j + \sum_{i=1}^{q}\beta_i \approx 1$, implying that the model is highly persistent. When this sum is equal to one current information remains relevant for all future forecasts, that is, the model is integrated. The FIGARCH model adjusts for integration by introducing a difference operator, termed d where $0 \le d \le 1$, into the conditional variance equation. Finally, although this study has considered only univariant models, it is possible to extend the GARCH framework to include multiple exogenous variables, termed as a GARCH-X model. This can be achieved by including another variable, say X_t, into the conditional variance equation.

In addition to using more advanced techniques to estimate volatility it will also be interesting to broaden the scope of the data set used in this study. The results in this paper only hold true for Microsoft equity, thus, the results are not robust enough to make any general conclusions regarding the validity of GARCH processes. Including more securities so that the entire market is represented will lead to more convincing and vigorous conclusions.

References

Berndt, E.K., B.H. Hall, R.E. Hall and J.A. Hausman. 1974. “Estimation Inference in Nonlinear Structural Models.” *Annals of Economic and Social Measurement*, 31, 307-27.

Black F. and Scholes M. 1973. “The Pricing of Options and Coorporate Liabilities.” *Journal of Political Economy*, 81, 637-659.

Bollerslev, Tim. 1986. “Generalized Autoregressive Conditional Heteroskedasticity.” *Journal of Econometrics*, 31, 307-27.

Bollerslev, Tim, Ray Y. Chou, and Kenneth F. Kroner. 1992. “ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence.” *Journal of Econometrics*, 52, 5-59.

Engle, Robert F. 1982. “Autoregressive Conditional Heteroscedasticity with Estimates of Variance of United Kingdom Inflation.” *Journal of Applied Econometrics*, 4, 1-21.

Itô K. 1951, “On Stochastic Differential Equations,” *Memoirs, American Mathematical Society*, 4, 1-51.

Merton M.C. 1973. “Theory of Rational Option Pricing.” *Bell Journal of Economics and Management Science*, 4, 141-83.

Nelson, Daniel B. and Charles Q. Cao. 1992. “Inequality Constraints in the Univariate GARCH Model,” *Journal of Business and Economic Statistics*, 45, 267-90.