

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

# COMPARING GARCH MODELS FOR GOLD PRICE DATA, USING A STATISTICAL LOSS FUNCTION APPROACH AND AN OPTION PRICING APPROACH

BLAKE CUNINGHAM

## *Abstract*

Derivative instruments that rely on the price of gold are traded in large volumes. A significant number of these instruments are influenced by the volatility of gold price movements. Hence, it is important to understand the volatility of this commodity when developing successful trading and hedging strategies. In this thesis, use is made of various GARCH models that are evaluated using both in-sample and out-of-sample criteria. Thereafter, a basic options pricing straddle strategy is developed, to evaluate the performance of these models in pecuniary terms. The GARCH models that differentiate between positive and negative effects (skewness) outperformed those that made use of a symmetrical distribution in terms of both in-sample and out-of-sample tests; whilst the use of different distributions for the mean adjusted returns produced mixed results. When these models are incorporated in a basic options strategy it is shown that it is possible to derive a theoretical profit/loss for one of the parties, where the value of the theoretical profit/loss depends on the structure of the model that is employed.

## 1. INTRODUCTION

During the past year, when many financial assets have provided below average returns (as compared to recent decades), the price of gold has increased significantly, as many investors seek refuge from the effects of the recent global financial crisis.<sup>1</sup> This has resulted in significant changes to the composition of the gold market, as the demand for the commodity by investors has recently surpassed that demanded by jewellery manufacturers.<sup>2</sup> This change to the market has also impacted on the volatility of this commodity, as price movements are now largely influenced by speculative investors, who tend to trade more frequently. Understanding these changes in the volatility of this commodity is of great importance to traders, particularly those who are making use of derivative instruments (that are largely based on the volatility of the underlying asset).

To further our understanding of the volatility of gold price movements, this thesis makes use of various general autoregressive conditional heteroskedasticity (GARCH) models. Before fitting these models, the autocorrelation structure of gold

---

<sup>1</sup> Between 1 January 2010 and 31 December 2010 the price of gold rose by 23.6% (in US dollar terms), whilst the MSCI World Index rose by 9.1%.

<sup>2</sup> The Erste Group (2010:38) state that the current demand for gold by investors is 37%, compared to 4.7% in 2000.

price data and the distribution of the returns are investigated. It is noted that whilst the returns appear to be stationary, there is evidence of autocorrelation in the variance. In addition, the data also appears to have a leptokurtic distribution. Respective GARCH, EGARCH and GJR-GARCH models are then used to describe the volatility structure in the data.<sup>3</sup> Each of these models is implemented using both normal distributions and student's  $t$  distributions before they are evaluated using rigorous in-sample procedures. Thereafter, extensive out-of-sample evaluations are performed, using various loss functions.

As part of a final evaluation exercise, use is made of the GARCH models in a basic options pricing strategy to determine which of the models is able to generate the greatest profit margin for each of the respective investors (buy and sell side). This strategy makes use of a basic Black-Scholes framework for the derivation of a straddle trading strategy that is applied over short horizons. Hence, this investigation effectively seeks to determine whether the respective GARCH models are able to effectively price for the probability of large changes in the price of gold.

In the following section, the literature that is relevant to this investigation is reviewed. Section 3 describes the theoretical framework and section 4 describes the data. The programming code that is used to model the data is described in section 5 and a discussion of the results is contained in section 6. Section 7 describes the framework of the options strategy as well as the respective results. Section 8 concludes.

## 2. LITERATURE REVIEW

The theoretical and technical groundwork for this thesis follows in subsequent sections. This section is just a brief review of the literature concerning the effectiveness of forecasting volatility with GARCH models.

Many papers report the persistence of volatility<sup>4</sup>, demonstrating that financial time series do indeed have model able volatility characteristics. However, the findings and opinions regarding the forecasting results were often sceptical about the ability of the GARCH models (Bollerslev and Anderson, 1998). These concerns primarily related to the GARCH models being able to explain very little of the variability in the *ex-post* squared returns. Bollerslev and Anderson (1998: 886) made the point that comparing volatility forecasts to the *ex-post* daily squared returns is not a good way to measure forecasting accuracy because of the noise associated with the measure. They suggested that higher frequency data was better, in the form of cumulative squared returns, and that the forecasting results using this measure were

---

<sup>3</sup> The GARCH model was originally developed by Bollerslev (1986), the exponential GARCH (or EGARCH) was developed by Nelson (1991), and the Glosten-Jagannathan-Runkle GARCH (or GJR-GARCH) was developed by Glosten, Jagannathan and Runkle (1993).

<sup>4</sup> Bollerslev, Chou and Kroner (1992); Bollerslev, Engle and Nelson (1994); Ghysels, Harvey and Renault (1996); Shephard (1996) all provide surveys of results pertaining to the persistency of volatility in time series data.

notably better and significant. The higher the frequency, the better the measure was, to the point that infinite frequency (infinitesimally small intervals between observations) actually produced a measure of genuine volatility. Hsieh (1991) and Fung and Hsieh (1991) confirm this finding as they reported that using 15 minute intervals yielded significantly higher fit of volatility as compared to inter-day measures. Bollerslev and Anderson (1998: 900) found that by increasing their intervals from an initial daily observation, to an hourly total of 24 observations per day with their currency data, the  $R^2$  statistic improved from 0.047 to 0.331. This thesis therefore makes good use of intra-day observations in its methodology.

Bollerslev and Anderson (1998: 901) ultimately concluded that GARCH models did provide good volatility forecasts.

In later works, the literature tested which particular GARCH models were the best. It was found that this often depended on the particular data – Hanson and Lunde (2005) found that for DEM/USD the basic GARCH(1,1) was not outperformed by any other more complex model, whilst there was a drastic improvement on this model by leverage models for IBM stock price data. Christoffersen and Jacobs (2004) found that leverage models performed better in an option pricing framework – a result that appears to disagree with the findings of this thesis.

Marcucci (2005) found that leverage models using fatter tailed distributions performed best for S&P index data. The literature reviewed therefore seems to agree that GARCH models are good at forecasting cumulative returns of high frequency data, where often leverage models outperform symmetric models, and fatter-tailed distributions incorporated into forecasts added to forecasting success.

No papers were found that tested different GARCH models for gold price data, and none were found that contrasted statistical loss function results and option pricing results. This paper therefore attempts to add to the literature by way of studying volatility for gold, and the differing results between accuracy checking methods.

### 3. THEORETICAL FRAMEWORK

The daily closing gold price per ounce was considered ( $P_t$ ) and the corresponding rate of log returns  $r_t$  was produced, which is defined as the continuously compounded rate of return

$$r_t = 100[\log(P_t) - \log(P_{t-1})] \quad (1)$$

where the index  $t$  represents daily closing observations and  $t = -R + 1, \dots, n$ . Within this, the sample consists of an estimation (in-sample) period denoted

$t = -R + 1, \dots, 0$  and an assessment (out-of-sample) period denoted  $t = 1, \dots, n$  which have  $R$  and  $n$  observations respectively<sup>5</sup>.

Following Bollerslev<sup>6</sup> (1986: 309), a GARCH(1,1) model for the series of returns on the gold price can be written as

$$r_t = \mu + \varepsilon_t \quad (2)$$

$$\varepsilon_t = z_t \sqrt{h_t} \quad (3)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta h_{t-1} \quad (4)$$

where  $\alpha_0 > 0, \alpha_1 > 0$  and  $\beta > 0$  to ensure positive conditional variance and the disturbance  $z_t$  is an *i.i.d.* process with zero mean and unit variance.

To alleviate the concern relating to non-negativity constraints on the ARCH parameters, Nelson (1991: 349) proposed the exponential GARCH, or EGARCH model. This model, with the conditional variance logged, avoids the parameter restrictions to ensure positive conditional variance and can be written as

$$\log(h_t) = \alpha_0 + \alpha_1 \left| \frac{\varepsilon_{t-1}}{h_{t-1}} \right| + \xi \frac{\varepsilon_{t-1}}{h_{t-1}} + \beta_1 \log(h_{t-1}) \quad (5)$$

This model would also account for leverage effects, where positive and negative returns affect volatility in different ways. That is, it takes into account the tendency for volatility to rise when returns fall and to fall when returns rise<sup>7</sup>.

Further modifications to the standard GARCH model include those of Glosten, Jagannathan and Runkle (1993) with their GARCH-GJR. The model incorporates the asymmetry between positive and negative returns in that it allows the conditional variance to respond differently to them. The model is described as

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 [1 - I_{\{\varepsilon_{t-1} > 0\}}] + \xi \varepsilon_{t-1}^2 I_{\{\varepsilon_{t-1} > 0\}} + \beta_1 h_{t-1} \quad (6)$$

where  $I_{\{\omega\}}$  is a dummy variable equal to one if  $\omega$  is positive, and zero otherwise. Therefore, the effect of a positive shock on the conditional variance is given by  $\xi \varepsilon_{t-1}^2$ , whilst a negative shock has an effect equal to  $\alpha_1 \varepsilon_{t-1}^2$ . Obviously if  $\alpha_1 > \xi$ , negative shocks will have a greater impact on the conditional variance and a leverage effect will be present<sup>8</sup>.

<sup>5</sup>  $R$  is the number of observations in-sample, and  $n$  is the number of observations out-of-sample.

<sup>6</sup> Robert Engle (1982) was the first to propose ARCH modelling, but Benoit Mandelbrot (1963) was the first to note that volatility tended to cluster, with large deviations followed by large deviations, and small deviations followed by small deviations – regardless of sign.

<sup>7</sup> This was first noted by Fisher Black (1976).

<sup>8</sup> Conversely if  $\alpha_1 < \xi$  a leverage effect is also present, but of opposite nature. This relationship is not expected.

An increasingly popular alternative to the GARCH models is the stochastic volatility model. It may be specified with the same mean regression representation as above. However, the conditional variance of the model may be given by

$$\log(h_t) = \alpha_0 + \alpha_1 \log(h_{t-1}) + v_t \quad (7)$$

where  $v_t \sim NID(\delta, \sigma^2)$  is a random variable and  $N$  denotes a normal distribution. The above conditional variance equation contains no moving average components and is referred to as an SV(1,0) model. In a GARCH model, the conditional volatility is perfectly explained by past observations, whilst in a stochastic volatility model, additional uncertainty in volatility is allowed for by the introduction of a stochastic error term. The inclusion of this term may be useful in the case of gold's conditional volatility, which often appears to incorporate an element of randomness. However, the model was not used for the purposes of this thesis, as it has been noted to be poorer at forecasting than the simpler GARCH type models. The stochastic volatility model could however be useful in forecasting forward more realistic confidence intervals due to the additional stochastic element which would be prominent in Monte Carlo simulations. This thesis is however more interested in likely specific values for volatility, and so remains concerned with the GARCH family.

As is noted by Marcucci (2005: 4), a common finding in the GARCH literature is the leptokurtosis of the distributions of financial returns. The situation is no different for returns on gold (see table 1). As a result, it becomes appropriate to use fatter-tailed error distributions to take account of this. The above GARCH models are therefore run with both Gaussian and Student's  $t$  errors<sup>9</sup>.

## 4. INVESTIGATION OF THE DATA

### 4.1 General Data Description

The data employed in this paper are the daily closing gold prices<sup>10</sup> from 02 January 2002 to 26 November 2010 with 2318 observations. After conversion to a return series, one data point is lost and thus there are 2317 observations.

In addition to this, hourly prices were used to compute daily sums of hourly absolute returns as a proxy for the true volatility of the gold price over the out-of-sample period. This is in line with Marcucci (2005: 9) as it is a better measure for volatility than the traditional squared returns. Further comments on this will be made in the next section.

---

<sup>9</sup> The Generalised Error Distribution (GED) is also often used, and is used by Marcucci (2005), but was not used in this thesis due to computer software limitations.

<sup>10</sup> XAU:USD is the currency code for gold, which is the spot price for 24kt gold. The ETF (Exchange Traded Fund) for gold has the code GLD, and is pegged at one tenth of XAU:USD. The actual gold spot price data is the one used during the estimations.

The data for the gold price, both hourly and daily, was obtained from the website <http://www.fxhistoricaldata.com/XAUUSD/>.

Table 1 below shows some of the descriptive statistics for the returns on gold, resized into percentages. The mean of the series is slightly positive at just under 0.07% per day, with annualized yields of 17.4% nominal return and a 19.0% effective rate of return<sup>11</sup>. These returns are substantially positive, and indicate a strong and consistent upward trend in the price of gold, but should not be considered proof that the price of gold will always behave in this manner, in spite of the strongly statistically significant trend based on regression analysis<sup>12</sup>. The standard deviation of just over 1.2% indicates that the average movement above or below the mean of 0.07 is 1.2% - it is within this range (positive 1.28% and negative 1.14%) that most observations lie. However, the maximum deviation from the mean that has occurred ( $10.357 - 0.069 = 10.289\%$ ) shows that large deviations have occurred within the sample.

Given the above characteristics, it would be inappropriate to assume that returns are normally distributed as the largest deviation, which has an 8.47 sigma, should only occur once in 12 405 916 747 260 300 (over 12 quadrillion) periods. This distribution assumption is therefore quite unrealistic, as although most of the observations still fall within one sigma, larger deviations are trillions of times more likely to occur than would be expected by the normal distribution.

The Jarque-Bera test also confirms the non-normality of the returns series with a test statistic of 2515 against a critical value of 5.967, significantly rejecting the null hypothesis of normality. Furthermore, it is noted that the distribution has very fat tails, as the kurtosis is large. This implies that the use of the fatter-tailed student's  $t$  may be more appropriate, although a distribution with even fatter tails may be more appropriate. The skewness of the returns is found to be negative, indicating that negative returns are more likely to be further from the mean than positive returns.

Importantly for this paper, the LM(12) and  $Q^2(12)$ <sup>13</sup> statistics indicate the presence of significant ARCH effects up to 12 lags.

---

<sup>11</sup> The nominal rate is merely the mean multiplied by 252, whilst the effective rate is  $(1 + \text{meanreturn})^{252} - 1$ .

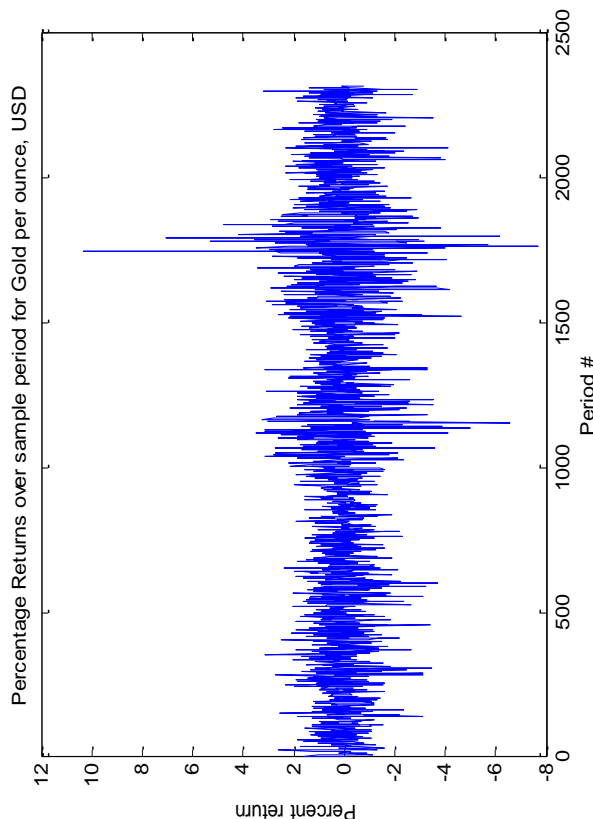
<sup>12</sup> This is merely the problem of induction (Popper, 1972) – no amount of highly significant positive regressions can imply that this will always be the case.

<sup>13</sup> LM(12) is the Lagrange Multiplier test for ARCH effects in the OLS residuals from the regression of the returns on a constant, while  $Q^2(12)$  is the corresponding Ljung-Box statistic on the squared standardised residuals (Marcucci, 2005: 9).

Table 1: Descriptive Statistics for Percentage Returns<sup>14</sup>

		Descriptive Statistics									
Mean	Standard Deviation	Minimum	Maximum	Skewness	Kurtosis	Normality Test	Crit (5%)	ARCH Test	Q(12)	Crit (5%)	
0,069		1,213	-7,707	10,357	-0,233	8,083	2515,578	5,967	165,281	25,685	21,026

Figure 1: Daily percentage returns of the gold price



<sup>14</sup> Note: The sample period is from 2 January 2002 to 26 November 2010 and includes 2317 observations. The normality test above is the Jarque-Bera test which follows a  $\chi^2$  distribution with 2 degrees of freedom under the null hypothesis of normally distributed errors. The LM(12) statistic is the statistic for the ARCH LM test on the residuals of the conditional mean regression up to 12 lags. This is performed under the null hypothesis of no ARCH effects and follows a  $\chi^2(q)$  distribution where q is the number of lags. The  $Q^2(12)$  statistic is the Lung-Box test on the squared residuals from the OLS conditional mean regression up to the 12<sup>th</sup> order. This statistic also follows a  $\chi^2(q)$  distribution where q is the number of lags. The 5% critical value for both tests is therefore 21.03. The standard errors for the skewness and kurtosis are  $\sqrt{24/T}=0.0509$  and  $\sqrt{24/T}=0.1018$  respectively, where T of course equals 2317.

Figure 2: Kernel Density of gold returns

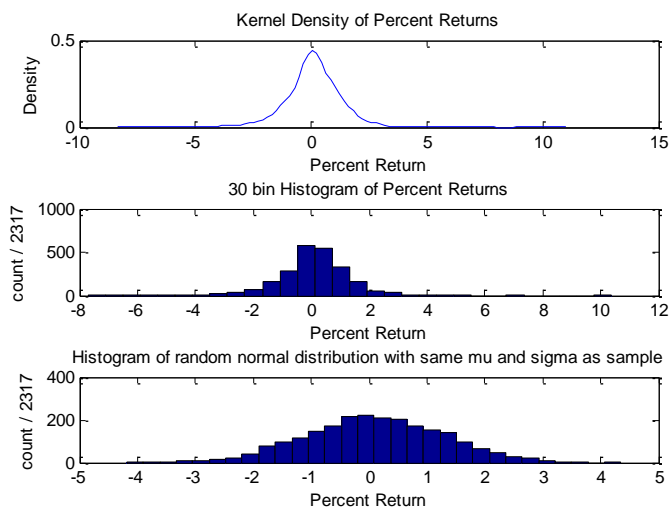


Figure 1 displays the returns on gold and the high volatility of the series over the entire sample. Most notably, the graph shows strong signs of volatility clustering. Unsurprisingly, the period exhibiting the highest volatility occurs after the 2008 financial crisis. One can also see the high volatility just before the end of the series (November 26, 2010) where the price of gold famously, but briefly, broke the \$1400 per ounce mark on the 9<sup>th</sup> of November. ACFs and PACFs of the returns data (figures 3 and 4, respectively) show that there is no need to include any correlation structure in the conditional mean equation, as the lags are not significant. The ACF (figure 5) of the squared returns, however, illustrates the relationship between returns and its lags through its second moment, which makes the case for some form of volatility modelling. The ACF and PACF of the intra-day volatility (figures 7 and 8, respectively) show significant signs of ARCH effects.

Figure 2 shows the kernel density of the gold returns. Notably, the majority of percentage daily returns are close to zero, and one can see the stark difference in shape to the normal distribution with similar moments.

Figure 3: Autocorrelation of returns

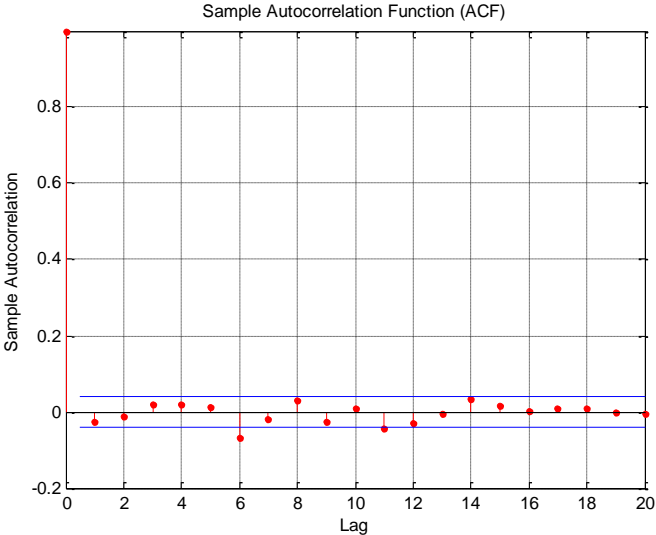


Figure 4: Partial autocorrelation of returns

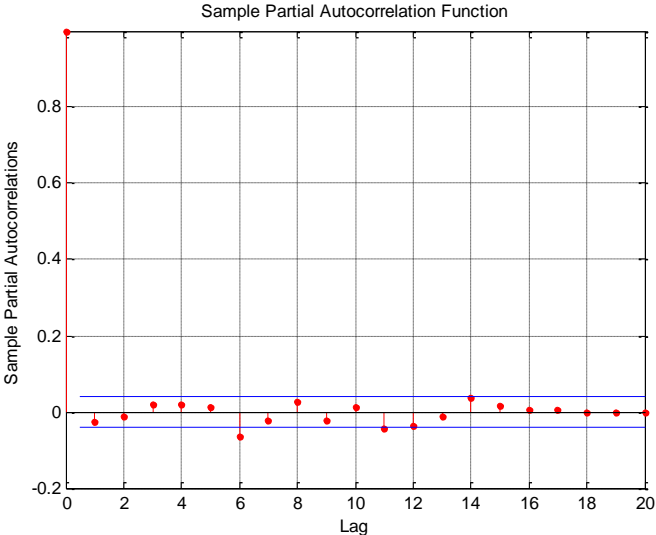


Figure 5: Autocorrelation of squared returns

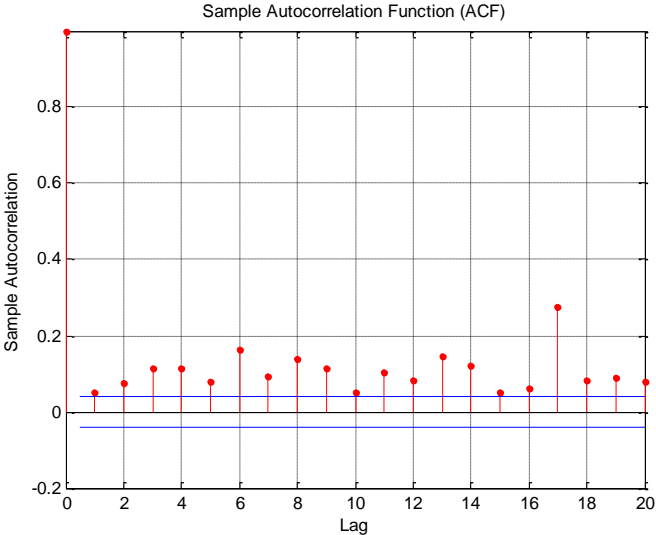


Figure 6: Partial autocorrelation of squared returns

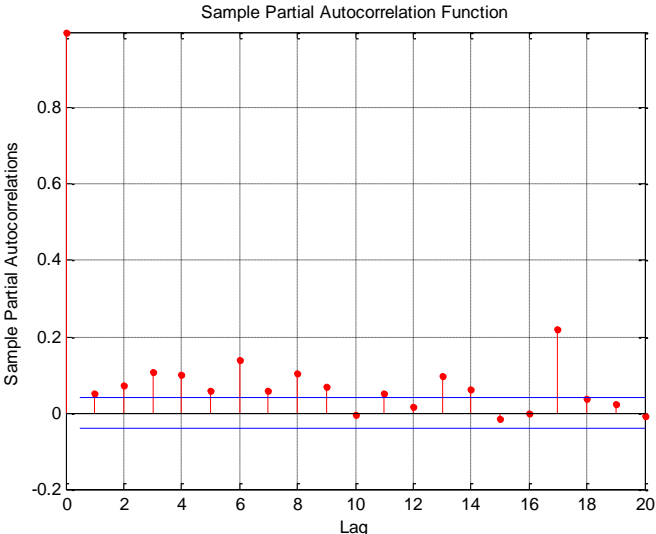


Figure 7: Autocorrelation of intra-day volatility

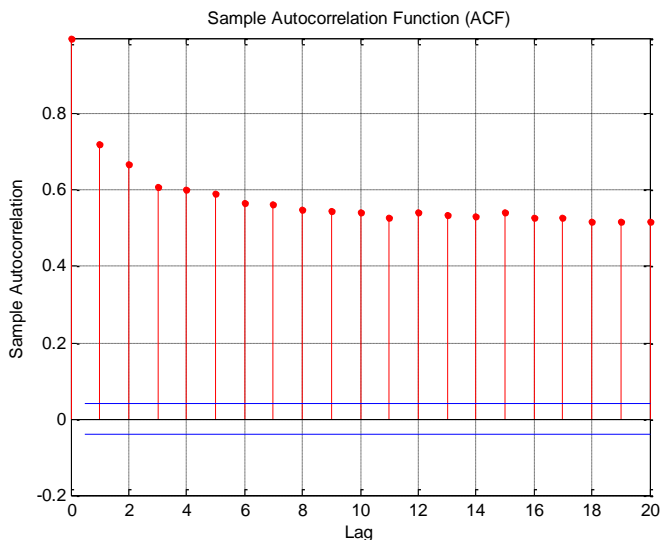
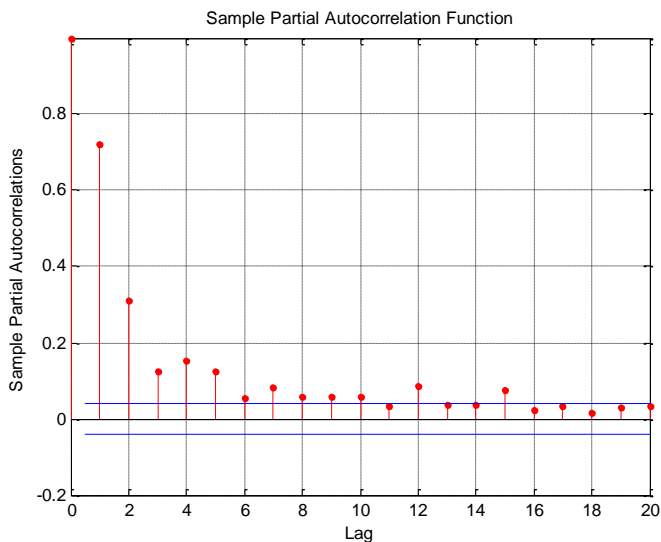


Figure 8: Partial autocorrelation of intra-day volatility



#### 4.2 Pareto Properties of Gold Returns<sup>15</sup>

In figure 2 it was noted that the returns series did not exhibit the characteristics of a normal distribution, where fat tails, or leptokurtosis, was observed. It therefore makes sense to see if the returns fit other distributions. Examples of other types of

<sup>15</sup> This section deals with the shape of the returns distribution, where the second moment is associated with volatility. The standard deviations of returns around a mean are distributed in the same way as the returns themselves, and standard deviation (as the square root of variance) is related to the second moment. This section describes how the distribution of the data is shaped relative to distributions that have fatter tails.

distributions include those from the Power law family, which model much higher probabilities for extreme observations.<sup>16</sup> One of the distributions in the Power law family includes the Pareto distribution, which was tested against the data. In order to obtain a visual fit of a Pareto distribution to demonstrate its similarity to the data, the moments were matched using the Method of Moments. Below is the Generalised Pareto Distribution (GPD)<sup>17</sup>:

$$f(x) = \frac{1}{\gamma} \left( 1 + k \frac{(x-\theta)}{\gamma} \right)^{-\left(\frac{1}{k}+1\right)} \quad (8)$$

where  $\theta$  is known as the location parameter,  $\gamma$  is the scale parameter, and  $k$  is the shape parameter. Together these parameters define the PDF of the generalised Pareto, and so finding their values is of importance in reproducing a Pareto similar to that of the gold returns. As we only know two moments of the absolute value of returns, the standard deviation and the mean, it is required that one of the Pareto parameters be done away with. This is easily done for  $\theta$  as we know that the lower limit of the one tailed distribution<sup>18</sup> is zero, and this defines the position of the PDF. Hence it is only required to find  $\gamma$  and  $k$ . The moments of the GPD are as follows:

Mean:

$$\bar{x}(\text{mean}) = \theta + \frac{\gamma}{1-k} \quad (9)$$

Variance:

$$\hat{\sigma}^2(\text{variance}) = \frac{\gamma^2}{(1-k)^2(1-2k)} \quad (10)$$

And, as stated, it is assumed that:

$$\theta = 0 \quad (11)$$

Therefore, rearranging equation x:

$$\gamma = \bar{x}(1 - k) \quad (12)$$

And solving simultaneously yields:

<sup>16</sup> This can be easily tested if one were to input the same large deviation into a Pareto PDF and a Normal PDF.

<sup>17</sup> This is the PDF used by the MATLAB software package.

<sup>18</sup> Because the GPD is a one-tailed distribution, and we are concerned about the absolute deviations of the data, the absolute values of the returns series was used for this exercise.

$$k = \frac{1}{2} - \frac{\bar{x}^2}{2\hat{\sigma}^2} \quad (13)$$

$$\gamma = \frac{\bar{x}}{2} + \frac{\bar{x}^3}{2\hat{\sigma}^2} \quad (14)$$

Hence, using the 2317 absolute percentage returns to calculate  $\bar{x}$  (0.8740) and  $\hat{\sigma}^2$  (0.8444)<sup>19</sup>, these values are next inputted into equations 13 and 14 to get values of:

$$k = 0.0477 \text{ and } \gamma = 0.8323$$

It is with these values that a random draw of 2317 i.i.d. numbers between 0 and 1 is generated by a random number generator<sup>20</sup> which can be applied to an inverse GPD cumulative distribution function in order to reproduce observations similar to the actual data. Figure 9 illustrates that the returns exhibit the same shape as the GPD, with the extreme values being within the same range, although the actual returns still exhibited a larger deviation than this particular random draw. The power law distributions are usually used to specifically model extreme values (Hosking and Wallis, 1987: 339), and may actually be more demonstrative of the return series if done in this fashion – i.e. only modelling the returns greater than one or two standard deviations. The main point though, is that the Pareto distribution needs to be considered in terms of allowing for the inordinately greater possibility of large deviations when compared to the usual normal distribution family<sup>21,22</sup>.

---

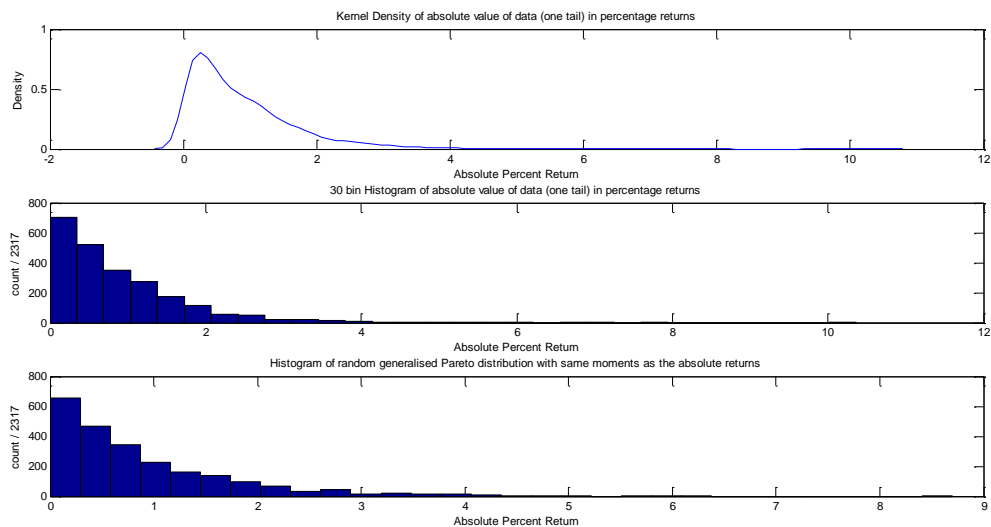
<sup>19</sup> These moments are, of course, different from the moments of the actual returns series because they are generated from the absolute data.

<sup>20</sup> The command `rand(2317,1)` produces a randomly uniform distribution between 0 and 1 of 2317 points.

<sup>21</sup> The Pareto distribution is often associated with the 80-20 principal when Pareto himself originally developed the Pareto distribution and showed that 80% of the wealth was owned by 20% of the populations. The gold return data is less extreme – 29% of the data accounts for 71% of the returns. This figure may not be shocking, but gets more extreme if one only models larger deviations.

<sup>22</sup> This applies specifically to the expectation of absolute deviations, as the Pareto is one-tailed.

Figure 9: Comparison of absolute returns with a Generalised Pareto Distribution



### 4.3 What is being forecasted

To compare the out-of-sample forecasting performance of competing models a measure of future volatility is needed. Marcucci (2005: 10) notes that in the literature, many researchers have used either the ex-ante or ex-post squared returns as a measure of future volatility. He argues that this measure represents a very noisy estimate of future volatility and can lead to incorrect conclusions regarding the real ability of GARCH type models when forecasting volatility. This thesis, in line with Marcucci, used intra-day variance calculated as the sum of the absolute log returns between hours during the day, where there are 24 hourly observations per day, and 2317 days. This method is preferred to using the independent variance of days because it is scale invariant – variance does not increase as the price level increases. The volatility forecasts of the models at the different horizons are denoted as  $\hat{h}_{t+h|t}$  and the actual or realised volatility as  $\hat{\sigma}_{t+1|t}^2$ .

Importantly, when comparing the forecasts, the volatility is converted back into terms of returns<sup>23</sup>. From the standard GARCH model<sup>24</sup>:

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta h_{t-1} \quad (15)$$

The returns are stationary about a mean:

<sup>23</sup> Although unorthodox, this was done as opposed to using the square of the mean-adjusted returns, because it was deemed to be easier to understand. In any case, there is no impact on the results.

<sup>24</sup> The derivation is no different for the EGARCH or GJR.

$$r_t = \mu + \varepsilon_t \quad (16)$$

The error on the returns series has the following structure:

$$\varepsilon_t = z_t \sqrt{h_t} \quad (17)$$

Where:

$$z_t = \frac{r_t - \mu}{\sqrt{h_t}} \quad (18)$$

Therefore:

$$r_t = \mu + z_t \sqrt{h_t} \quad (19)$$

Where  $z$  is distributed with zero mean and a variance of one:

$$z_t \sim D(0,1) \quad (20)$$

Squaring both sides of equation 17 yields:

$$\varepsilon_t^2 = h_t z_t^2 \quad (21)$$

As  $z$  has a variance of one and a mean of zero, it follows that we can expect:

$$E_t(z_{t+i}^2) = 1 \quad (22)$$

Hence, the expectation of equation 21:

$$E_t(\varepsilon_{t+i}^2) = E_t(h_{t+i}) \quad (23)$$

And the root of the expectation (given that  $h$  is positive):

$$\sqrt{E_t(\varepsilon_{t+i}^2)} = \sqrt{E_t(h_{t+i})} = E_t(\sqrt{h_{t+i}}) \quad (24)$$

And as the expectation of returns is:

$$E_t(r_{t+i}) = E_t(\mu_{t+i}) + E_t(\sqrt{h_{t+i}}) \quad (25)$$

Hence, absolute realised returns are compared against a forecast of the return series' mean component plus the root of the volatility forecast. From this point onwards the following notation applies:

$$r_{t+i} = \hat{\sigma}_{t+1|t}^2 \quad (26)$$

And

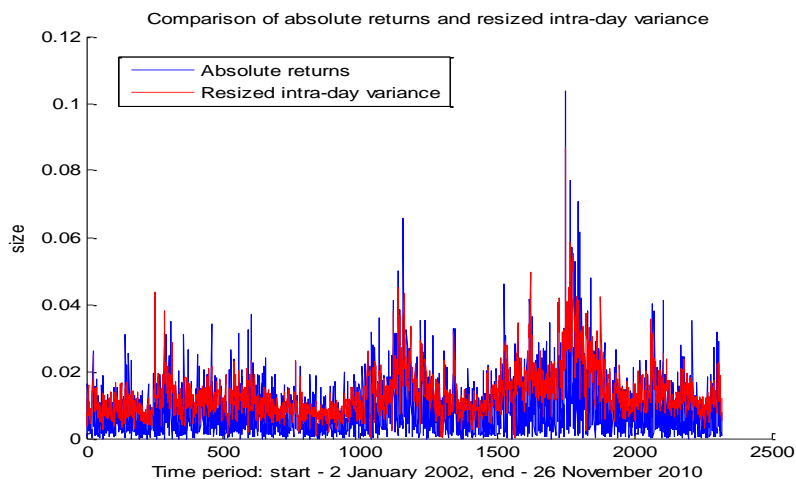
$$E_t(r_{t+i}) = \hat{h}_{t+h|t}, \quad (27)$$

The intra-day variance then needs to be resized to the return series. This is done by means of a search algorithm<sup>25</sup> that seeks to identify that value of  $\alpha$  that minimises the sum of least square errors from the expression:

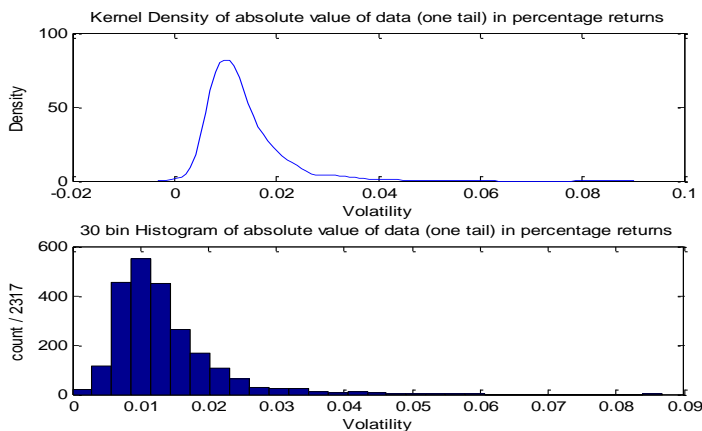
$$(vol_t - \alpha|r_t|)^2 \quad (28)$$

This  $\alpha$  was then used to divide the intra-day variance values by, so that they were the same size as the absolute returns. This means that they could then be compared to the forecast of the  $\mu$  plus  $\sqrt{h_t}$ . The alpha that was found to be optimal was 2.705781. It is noted that on visual inspection (figure 10), the intra-day data seems more stable, and has a thicker base. This is because returns are subject to losing substantial information if a day's opening and closing prices are close together, in spite of large fluctuations within a particular day. This can also be seen with the peak of the density diagram being noticeably different from zero, as compared to the density plots of the absolute returns - see figure 11.

Figure 10: Comparison of resized intra-day volatility with absolute returns



<sup>25</sup> The Solve add-on in MS Excel easily allows one to minimise a sum of squares by changing selected parameters.

*Figure 11: Kernel density and histogram of intra-day volatility*

## 5. EXPLANATION OF THE EXPERIMENT CODE

### 5.1 Graphical Description of the Procedure

Prior to a discussion of the statistical loss functions and the associated results, it is necessary to cover the procedure that was programmed in MATLAB as this provides for a more adequate grasp of the results tables. The methodology can be explained using a brief series of diagrams. The out-of-sample fit measures the accuracy of the forecasts of the different models against data of actual intra-day volatility (described previously), and actual inter-day absolute returns. In order to produce a matrix of forecast data for a particular model and a particular horizon, actual data before the forecast period is used to get parameter values for the forecast function. In figure 12 the idea is made clear: the darker shaded cells represent the actual return data for the series that is used in parameter estimation; the parameters are then inputted into the forecast function which then forecasts a vector of values (the lighter shaded cells); those values are then slotted into the column (see figure 13) corresponding to the loop number. On the next iteration of the loop the amount of real observations<sup>26</sup> used to forecast the next horizon increases by one, and the horizon forecasted shifts forward by one period (regardless of horizon length). Hence, subsequent forecasts overlap in terms of time reference by the length of the horizon less one. In each loop the horizon of forecasts is placed into a column (figure 13) and stored for later analysis. There are 252 loops for every permutation of horizon and model (252 was chosen as there are 252 trading days in

<sup>26</sup> Note that no forecasted values are used in any parameter estimation. Each loop updates the data used to estimate the parameters by one period with the actual return for that day. Therefore, all forecasts of  $t+i$  are based on parameter estimation given our knowledge at  $t$ .

a year). There are six models and four horizons (1, 5, 10 and 22 days) - hence the loop produces 24 matrices<sup>27</sup>.

Figure 12: Loop logic

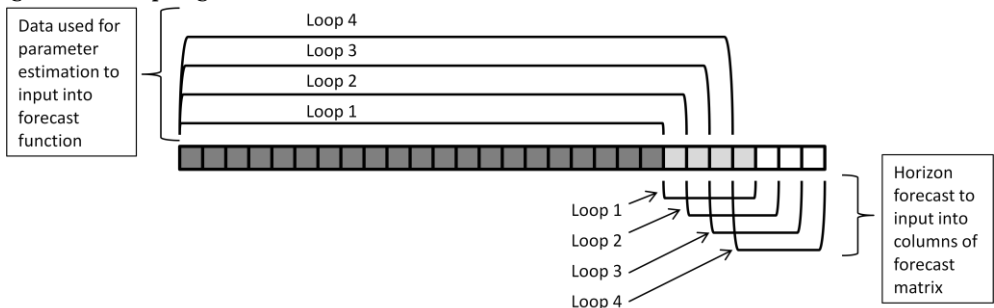
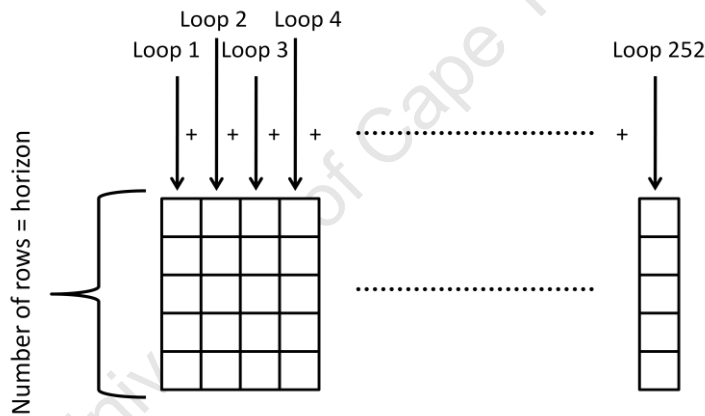


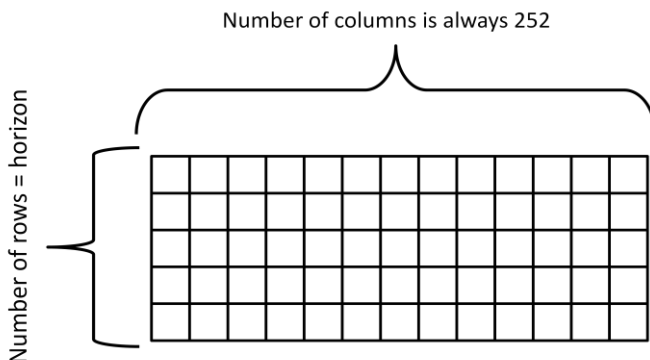
Figure 13: Formation of forecast matrix



Each matrix produced resembles the one in figure 14 – it will be 252 columns wide, and either 1, 5, 10 or 22 rows tall.

<sup>27</sup> This procedure could be shortened in order to use just one 22x252 matrix per loop per model, but the longer procedure used makes the coding simpler.

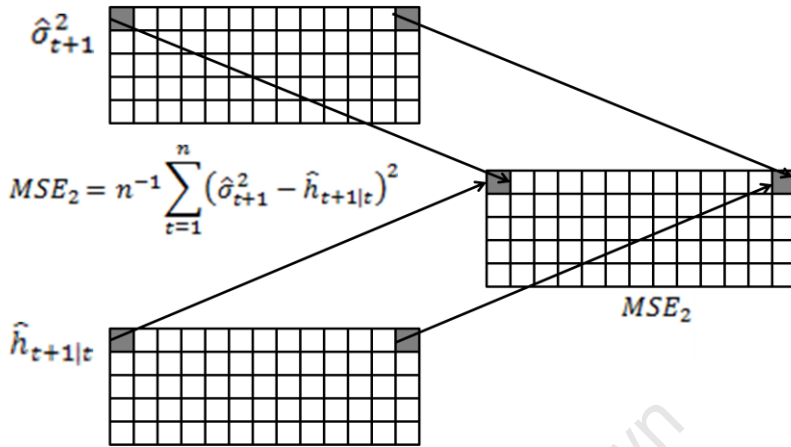
Figure 14: Completed forecast matrix



While the aforementioned looping process is running, 8 other matrices are being formed based on realised volatility. Four matrices each of intra-day and inter-day data corresponding to dimensions of  $1 \times 252$ ,  $5 \times 252$ ,  $10 \times 252$  and  $22 \times 252$ . It is these matrices that proxy for  $\hat{\sigma}^2$  in the calculation of the statistical loss functions (MSE, MAD etc.). When calculating the loss functions the process portrayed in figure 15 applies. In figure 15 the example of calculating an MSE2 is used: the top matrix is comprised of the realised volatility, whilst the bottom matrix is comprised of the forecasted volatility corresponding to the same time reference. Individual elements of the matrices for each time reference (row  $m$ , column  $n$ ) are then transformed using one of the loss functions (in this case the MSE2) and the result is placed into the corresponding reference in a third table of identical dimension. The individual  $\text{horizon\_length} \times 252$  elements of the third matrix are summed and then divided by  $n$  ( $\text{horizon\_length} \times 252$ ). This happens for all six models, times all four time horizons, times seven different statistical loss functions (less R2LOG for the squared returns), times two sets of realised volatility (inter-day and intra-day). All in all, we then get 312 results of interest.

In addition to this, another “model” is generated using the notion of Nassim Taleb’s taxi driver - this adds 44 more results for comparative purposes.

Figure 15: Calculation of the loss functions



### 5.2 Nassim Taleb's Taxi Driver

In his now infamous book “The Black Swan”, Nassim Taleb offers the following, debatable, comments about GARCH modelling:

*“The econometrician Robert Engle, an otherwise charming gentleman, invented a very complicated statistical method called GARCH and got a Nobel for it. No one tested it to see if it has any validity in real life. Simpler, less sexy methods fare exceedingly better, but they do not take you to Stockholm.”*

(Taleb, 2010: 156)<sup>28</sup>

In another part of the book Taleb introduces a hypothetical taxi driver, one who makes predictions about the future based on the last known observation<sup>29</sup>. Hence this taxi driver is compared against the GARCH models by forecasting a horizontal line equal in height to the last observed value of both absolute returns, and of intra-day volatility. To be fair, the context that Taleb uses this example is with regards to far more complicated, but more stable and long-term data such as GDP, or any general quarterly/yearly economic indicator – shocks in GDP, for example, are not as volatile as the data in this data set. However, the notion is used here as a sort of check that the GARCH models do improve forecasts over a very crude method. In addition, Taleb is indeed more concerned about the margin of error associated with

<sup>28</sup> Engle was, of course, also awarded the Nobel prize for his contribution to cointegration.

<sup>29</sup> Technically, this is a random walk model without drift.

prediction<sup>30</sup> as opposed to specific predictions, which he perhaps correctly writes off as being at best a waste of time, and at worst dangerous. But, as long as one needs a “most likely scenario” (expectation) and a specific number prediction, where a shock would not have disastrous consequences, then it is safe to attempt to predict these specific numbers.

## 6. EVALUATION OF VOLATILITY FORECASTS

### 6.1 Statistical Loss Functions

With the above theoretical framework and data description covered, I progress to compare the estimated models on the basis of their performance on seven statistical loss functions. Instead of choosing one loss function as the “best and unique,” use was made of the seven below. Each of these have different interpretations and collectively they provide a more complete evaluation of forecasting performance. The criteria in (29) and (30) below are the typical mean squared error metrics. The R2LOG (32) loss function has the feature of penalizing the volatility forecasts asymmetrically in low and high volatility periods. At a first glance, the QLIKE function (31) does not make much intuitive sense as it does not appear that all values of  $\hat{h}_{t+1|t}$  and  $\hat{\sigma}_{t+1}^2$  close to each other will yield lower values. This is because the log of a small variance will yield a very negative number. As a result, not as much emphasis is placed on this statistic but it is included for the sake of completeness. The MAD functions in (33) and (34) provide value in that they are more robust to the presence of outliers than the MSE criteria, and the HMSE is valuable in that it adjusts for heteroskedasticity. Further details of, and references to, these loss functions may be found in Marcucci (2005: 11).

$$MSE_1 = n^{-1} \sum_{t=1}^n (\hat{\sigma}_{t+1} - \hat{h}_{t+1|t}^{1/2})^2 \quad (29)$$

$$MSE_2 = n^{-1} \sum_{t=1}^n (\hat{\sigma}_{t+1}^2 - \hat{h}_{t+1|t})^2 \quad (30)$$

$$QLIKE = n^{-1} \sum_{t=1}^n (\log \hat{h}_{t+1|t} + \hat{\sigma}_{t+1}^2 \hat{h}_{t+1|t}^{-1}) \quad (31)$$

$$R2LOG = n^{-1} \sum_{t=1}^n [\log(\hat{\sigma}_{t+1}^2 \hat{h}_{t+1|t}^{-1})]^2 \quad (32)$$

$$MAD_1 = n^{-1} \sum_{t=1}^n \left| \hat{\sigma}_{t+1} - \hat{h}_{t+1|t}^{1/2} \right| \quad (33)$$

$$MAD_2 = n^{-1} \sum_{t=1}^n \left| \hat{\sigma}_{t+1}^2 - \hat{h}_{t+1|t} \right| \quad (34)$$

---

<sup>30</sup> How accurate predictions have been in the past, and how accurate they are likely to be in the future in spite of historical data – “unknown unknowns”.

$$HMSE = n^{-1} \sum_{t=1}^n (\hat{\sigma}_{t+1}^2 \hat{h}_{t+1|t}^{-1} - 1)^2 \quad (35)$$

Unfortunately, the R2LOG measure could not be used to compare the forecasts against the squared returns, as some of the inter-day closing prices remained unchanged, and hence forced a log of zero, which is mathematically undefined. In addition, when comparing the taxi driver forecasts, the QLIKE is also undefined due to the log of zero being present and the HMSE is undefined as it requires division by zero.

## 6.2 In-sample Results

The in-sample results concern the fit of the various models to the same data used in the model estimation. The data used was the full 2317 return observations. In addition, the models were compared to the corresponding 2317 intra-day volatility measurements. It is important to note that the model with the best in-sample fit is not necessarily the model with the best out of sample fit. All the models used the errors from a return series with a constant (C), and all these constants were statistically significant in estimation. The Akaike and Bayesian Information Criteria ranked the GARCH-GJR with student's  $t$  distribution as being the best fit – a strong result as these criteria value parsimony, and this is one of two models with the most parameters, 6. The GARCH-GJR-T also performed the best overall with the sum of its ranks being the lowest – however, a surprising result was that its more parsimonious version with just a normal distribution achieved better ranks amongst the statistical loss functions. The overall result that the symmetric “vanilla” GARCH models (with  $t$  and normal distributions) achieved the worst ranks (5<sup>th</sup> and 6<sup>th</sup>, respectively) was expected prior to running the estimations.

The results vary somewhat when comparing the fit against absolute returns, as opposed to the intra-day volatility. The GARCH-GJR-T model was not the best, and the three best ranked models were those that used a normal distribution. This result is surprising as we saw earlier that returns do not conform to the normal distribution. Even though we are interested in forecasting volatility, and so prefer the intra-day comparison, the traditional approach to this kind of work has been to compare the volatility models to return data (Marcucci, 2005: 9).

Table 2: In-sample goodness of fit with loss function results against intra-day volatility.

In-sample results with intra-day loss function fit																							
Model	#Par	AIC	Rank	BIC	Rank	Log(L)	Rank	MSE1	Rank	MSE2	Rank	QLIKE	Rank	MAD1	Rank	MAD2	Rank	R2LOG	Rank	HMSE	Rank		
GARCH - T	5	-14374,2	3	-14345,5	3	7192,1	3	0,0000275	3	0,0004252	4	-3,3792856	5	0,0035004	5	0,0151698	6	0,1676376	6	0,1430888	5		
GARCH - N	4	-14288,8	6	-14265,8	6	7148,4	6	0,0000276	4	0,0004218	3	-3,3792846	6	0,0034736	3	0,0150235	4	0,1638771	3	0,1499692	6		
EGARCH - T	6	-14380,3	2	-14345,8	2	7196,2	2	0,0000290	6	0,0004339	6	-3,3801080	4	0,0035181	6	0,0151214	5	0,1675474	5	0,1390266	3		
EGARCH - N	5	-14295,8	5	-14267,1	5	7152,9	5	0,0000286	5	0,0004270	5	-3,3805608	3	0,0034834	4	0,0149663	3	0,1640226	4	0,1423585	4		
GJR - T	6	-14386,5	1	-14352,0	1	7199,3	1	0,0000266	2	0,0004119	2	-3,3821994	2	0,0034326	2	0,0148364	2	0,1631232	2	0,1332933	1		
GJR - N	5	-14304,6	4	-14275,8	4	7157,3	4	0,0000266	1	0,0004074	1	-3,3824871	1	0,0034042	1	0,0146968	1	0,1599371	1	0,1372225	2		

Table 3: In-sample loss function results against absolute returns

In-sample inter-day fit												
Model	MSE 1	Rank	MSE 2	Rank	QLIKE	Rank	MAD 1	Rank	MAD 2	Rank	HMSE	Rank
GARCH - T	0,0000793	4	0,0023410	4	-3,7203547	5	0,0071761	4	0,0397151	4	0,4551003	3
GARCH - N	0,0000770	1	0,0022622	1	-3,7289619	1	0,0070217	1	0,0389969	1	0,4633719	6
EGARCH - T	0,0000801	5	0,0023685	6	-3,7182380	6	0,0072174	6	0,0398919	6	0,4512313	2
EGARCH - N	0,0000785	2	0,0023117	3	-3,7245782	3	0,0071089	3	0,0393825	3	0,4562366	5
GJR - T	0,0000805	6	0,0023570	5	-3,7207375	4	0,0071979	5	0,0397371	5	0,4507456	1
GJR - N	0,0000788	3	0,0023006	2	-3,7270030	2	0,0070888	2	0,0392259	2	0,4562234	4

The model parameters are all significant – the standard errors of all the parameters are reported in brackets beneath them in table 4. For the leveraged models, one can see statistically significant differences between ARCH and Leverage (with the sign even changing under GJR specification) suggesting that there certainly are leverage effects within the data. i.e. positive and negative shocks should have differently weighted effects on volatility estimates.

Table 4: Model Parameters<sup>31</sup>

Model Parameters						
Parameter	GARCH-T	GARCH-N	EGARCH-T	EGARCH-N	GJR-T	GJR-N
C	0,0009457	0,0006826	0,0010652	0,0008859	0,0010327	0,0008458
	0,0002039	0,0002091	0,0002028	0,0002146	0,0002033	0,0002131
$\omega$	0,0000009	0,0000010	-0,0426019	-0,0364838	0,0000006	0,0000005
	0,0000004	0,0000003	0,0227470	0,0169502	0,0000003	0,0000003
ARCH	0,0401598	0,0439187	0,0819669	0,0913146	0,0658521	0,0683229
	0,0077512	0,0054818	0,0149463	0,0098968	0,0129889	0,0088961
GARCH	0,9540410	0,9494506	0,9950756	0,9954631	0,9589784	0,9551957
	0,0088668	0,0069284	0,0025449	0,0019141	0,0080136	0,0057309
Leverage	-	-	0,0431669	0,0374057	-0,0496613	-0,0480394
	-	-	0,0107291	0,0072088	0,0139301	0,0098802
DoF	6,4720712	-	6,6648389	-	6,7574208	-
	0,9355622	-	0,9665775	-	0,9883047	-

### 6.3 Out-of-sample Results

The out of sample forecast results are this thesis' main concern. They give an indication on the forecasting potential for each of the models. The results seem to be in agreement with the in-sample results in terms of model rank. The GARCH-GJR-T achieved the best rank for the intra-day comparisons, across all time horizons. This was closely followed by the EGARCH-T, whilst the GARCH-N achieved the worst model rank. Taleb's taxi driver didn't fare well at all, performing

<sup>31</sup> C is the mean of the return series used to generate the errors. Omega is the long run mean, or the constant, in the GARCH models. ARCH is the autoregressive coefficient, whilst GARCH is the moving average coefficient of the GARCH models. Leverage is another autoregressive coefficient, but for negative shocks only, where it applies. DoF is the degrees of freedom for the  $t$  distribution, which controls the optimal kurtosis.

slightly worse than the GARCH-N. However, it was not the worst ranked every time – at the 1 day horizon using the MSE1 loss function, it ranked 5<sup>th</sup> beating out both GARCH models. This result seems less to do with the inability of GARCH to forecast, and perhaps more to do with GARCH not being the best model for the data.

The results as compared to the absolute returns show a similar picture. The leveraged models perform best, but surprisingly the models making use of the normal distribution all ranked higher – the complete opposite of the in-sample results. Once again, Taleb’s taxi driver performed the worst of the lot.

One can also see the trend of the average errors to be larger for longer horizons. This is important to note, as it demonstrates the deteriorating forecasting power of the models the further into the future they try to predict.

Table 5: 1-step-ahead intra-day out-of-sample results

1-step ahead volatility forecasts																
Model	MSE 1	Rank	MSE 2	Rank	QLIKE	Rank	MAD 1	Rank	MAD 2	Rank	R2LOG	Rank	HMSE	Rank	Rank Count	Final Rank
GARCH - T	0,000019240	6	0,000324888	5	-3,372103392	5	0,002973322	5	0,013042455	5	0,098829555	5	0,144610885	5	36	5
GARCH - N	0,000019696	7	0,000332012	6	-3,370048157	6	0,002990433	6	0,013123378	6	0,100250231	6	0,155420132	7	44	6
EGARCH - T	0,000018201	2	0,000303721	2	-3,376778670	2	0,002884557	2	0,012614836	2	0,091972003	3	0,125464086	2	15	2
EGARCH - N	0,000018932	4	0,000315403	4	-3,374001474	4	0,002918578	4	0,012763978	4	0,094621847	4	0,138684278	4	28	4
GJR - T	0,000017977	1	0,000298462	1	-3,377673543	1	0,002868353	1	0,012537256	1	0,089848528	1	0,124316964	1	7	1
GJR - N	0,000018383	3	0,000304846	3	-3,376044242	3	0,002892507	3	0,012645611	3	0,091181421	2	0,132526337	3	20	3
Taxi	0,000019080	5	0,000337942	7	-3,366776709	7	0,003157455	7	0,013973537	7	0,111832726	7	0,153206820	6	46	7

Table 6: 5-step-ahead intra-day out-of-sample results

5-step ahead volatility forecasts																
Model	MSE 1	Rank	MSE 2	Rank	QLIKE	Rank	MAD 1	Rank	MAD 2	Rank	R2LOG	Rank	HMSE	Rank	Rank Count	Final Rank
GARCH - T	0,000020443	5	0,000347040	5	-3,367983641	5	0,003076258	5	0,013512553	5	0,105738759	5	0,157354500	5	35	5
GARCH - N	0,000020936	6	0,000354735	6	-3,365730742	6	0,003086631	6	0,013564968	6	0,107308672	6	0,169337477	6	42	6
EGARCH - T	0,000018606	2	0,000311665	2	-3,375262363	2	0,002924267	2	0,012796834	2	0,094658695	2	0,129630522	1	13	2
EGARCH - N	0,000019410	4	0,000324802	4	-3,372219076	4	0,002962844	4	0,012971038	4	0,097821901	4	0,143584061	4	28	4
GJR - T	0,000018535	1	0,000309223	1	-3,375655157	1	0,002916848	1	0,012756049	1	0,093444606	1	0,130098029	2	8	1
GJR - N	0,000019055	3	0,000317485	3	-3,373644437	3	0,002939398	3	0,012857531	3	0,095237408	3	0,139902825	3	21	3
Taxi	0,000025629	7	0,000443851	7	-3,351666525	7	0,003700996	7	0,016144192	7	0,140406652	7	0,194584847	7	49	7

Table 7: 10-step-ahead intra-day out-of-sample results

10-step ahead volatility forecasts																
Model	MSE 1	Rank	MSE 2	Rank	QLIKE	Rank	MAD 1	Rank	MAD 2	Rank	R2LOG	Rank	HMSE	Rank	Rank Count	Final Rank
GARCH - T	0,000021278	5	0,000363534	5	-3,363655179	5	0,003177680	5	0,013974708	5	0,111286129	5	0,165164605	5	35	5
GARCH - N	0,000021773	6	0,000371272	6	-3,361305047	6	0,003181188	6	0,013998298	6	0,112884447	6	0,17784341	6	42	6
EGARCH - T	0,000019333	2	0,000326085	2	-3,371292453	2	0,003015168	2	0,013209755	2	0,099505221	2	0,136487444	1	13	2
EGARCH - N	0,000020064	4	0,000338315	4	-3,368390098	4	0,003047290	4	0,013357915	4	0,102579063	4	0,149670515	4	28	4
GJR - T	0,000019307	1	0,000324546	1	-3,371609464	1	0,003009337	1	0,013172664	1	0,098592053	1	0,136847109	2	8	1
GJR - N	0,000019857	3	0,000333288	3	-3,369425801	3	0,003026068	3	0,013247547	3	0,100515315	3	0,147580992	3	21	3
Taxi	0,000027862	7	0,000485534	7	-3,344392601	7	0,003875327	7	0,016914219	7	0,153782183	7	0,206627663	7	49	7

Table 8: 22-step-ahead intra-day out-of-sample results

22-step ahead volatility forecasts																
Model	MSE 1	Rank	MSE 2	Rank	QLIKE	Rank	MAD 1	Rank	MAD 2	Rank	R2LOG	Rank	HMSE	Rank	Rank Count	Final Rank
GARCH - T	0,000022593	5	0,000395041	5	-3,356787874	5	0,003367494	6	0,014861353	6	0,122803567	5	0,178458281	5	37	5
GARCH - N	0,000023121	6	0,000403490	6	-3,354095126	6	0,003363007	5	0,014853836	5	0,124673968	6	0,192945105	6	40	6
EGARCH - T	0,000020438	1	0,000353736	2	-3,364846464	2	0,003195360	1	0,014065169	2	0,109845569	2	0,149826343	2	12	2
EGARCH - N	0,000021065	3	0,000365210	4	-3,362117138	4	0,003229612	4	0,014229531	4	0,113139346	4	0,161341865	3	26	4
GJR - T	0,000020473	2	0,000353443	1	-3,365210112	1	0,003199765	2	0,014064456	1	0,109350427	1	0,148899192	1	9	1
GJR - N	0,000021105	4	0,000363635	3	-3,362541878	3	0,003210386	3	0,014117467	3	0,111705302	3	0,162078856	4	23	3
Taxi	0,000032034	7	0,000555173	7	-3,332142495	7	0,004193667	7	0,018232159	7	0,173274700	7	0,243008526	7	49	7

Table 9: 1-step ahead inter-day out-of-sample results

1-step ahead volatility forecasts														
Model	MSE 1	Rank	MSE 2	Rank	QLIKE	Rank	MAD 1	Rank	MAD 2	Rank	HMSE	Rank	Rank Count	Final Rank
GARCH - T	0,000066959	6	0,002319200	5	-3,806534007	6	0,006869918	6	0,040217020	6	0,502776111	5	34	6
GARCH - N	0,000064850	3	0,002241959	2	-3,815662974	2	0,006713134	3	0,039460876	2	0,512390422	6	18	3
EGARCH - T	0,000066146	5	0,002320351	6	-3,806967732	5	0,006844510	5	0,040147752	5	0,484418124	1	27	5
EGARCH - N	0,000064365	2	0,002250349	3	-3,814987845	3	0,006710611	2	0,039498849	3	0,495647615	4	17	2
GJR - T	0,000065448	4	0,002292017	4	-3,810391205	4	0,006789983	4	0,039878408	4	0,487302091	2	22	4
GJR - N	0,000064074	1	0,002237700	1	-3,816876782	1	0,006685936	1	0,039367742	1	0,495366586	3	8	1
Taxi	0,000119074	7	0,003575894	7	Error	7	0,008187230	7	0,049112789	7	Error	7	42	7

Table 10: 5-step ahead inter-day out-of-sample results

5-step ahead volatility forecasts														
Model	MSE 1	Rank	MSE 2	Rank	QLIKE	Rank	MAD 1	Rank	MAD 2	Rank	HMSE	Rank	Rank Count	Final Rank
GARCH - T	0,000066737	6	0,002314905	6	-3,807437135	6	0,006861283	6	0,040193560	6	0,500366250	5	35	6
GARCH - N	0,000064523	3	0,002234687	2	-3,816990512	2	0,006699704	3	0,039414066	2	0,508812077	6	18	3
EGARCH - T	0,000065722	5	0,002308747	5	-3,809081516	5	0,006824420	5	0,040063695	5	0,481227610	1	26	5
EGARCH - N	0,000064068	2	0,002243422	3	-3,816677157	3	0,006699194	2	0,039452507	3	0,490692779	3	16	2
GJR - T	0,000065322	4	0,002290508	4	-3,811260285	4	0,006791582	4	0,039895900	4	0,483999923	2	22	4
GJR - N	0,000063807	1	0,002231773	1	-3,818308705	1	0,006677027	1	0,039335444	1	0,491309016	4	9	1
Taxi	0,000102425	7	0,003010554	7	Error	7	0,007336964	7	0,043885779	7	Error	7	42	7

Table 11: 10-step ahead inter-day out-of-sample results

10-step ahead volatility forecasts														
Model	MSE 1	Rank	MSE 2	Rank	QLIKE	Rank	MAD 1	Rank	MAD 2	Rank	HMSE	Rank	Rank Count	Final Rank
GARCH - T	0,000067877	6	0,002334692	6	-3,803978894	6	0,006922567	6	0,040432357	6	0,506415746	5	35	6
GARCH - N	0,000065626	3	0,002253939	1	-3,813505664	2	0,006760413	2	0,039652081	2	0,514787786	6	16	2
EGARCH - T	0,000066661	5	0,002322473	5	-3,806443932	5	0,006872101	5	0,040238786	5	0,486948302	1	26	5
EGARCH - N	0,000065263	2	0,002266188	3	-3,812966525	3	0,006763071	3	0,039703644	3	0,495734122	3	17	3
GJR - T	0,000066625	4	0,002315521	4	-3,807380332	4	0,006861298	4	0,040171298	4	0,489322890	2	22	4
GJR - N	0,000065016	1	0,002254137	2	-3,814675350	1	0,006739313	1	0,039576161	1	0,496665203	4	10	1
Taxi	0,000100306	7	0,002951888	7	Error	7	0,007196938	7	0,043130558	7	Error	7	42	7

Table 12: 22-step ahead inter-day out-of-sample results

22-step ahead volatility forecasts														
Model	MSE 1	Rank	MSE 2	Rank	QLIKE	Rank	MAD 1	Rank	MAD 2	Rank	HMSE	Rank	Rank Count	Final Rank
GARCH - T	0,000069871	6	0,002371469	6	-3,793084243	6	0,007039632	6	0,040858486	6	0,517651414	5	35	6
GARCH - N	0,000067635	3	0,002292086	1	-3,802128619	2	0,006880931	2	0,040103291	2	0,527210302	6	16	2
EGARCH - T	0,000068057	4	0,002342099	4	-3,797672446	4	0,006955932	4	0,040511048	4	0,495834157	2	22	4
EGARCH - N	0,000067416	2	0,002310772	3	-3,801208129	3	0,006899082	3	0,040223107	3	0,504044404	3	17	3
GJR - T	0,000068821	5	0,002360574	5	-3,795969648	5	0,006995520	5	0,040672476	5	0,495713213	1	26	5
GJR - N	0,000067090	1	0,002295684	2	-3,803362367	1	0,006868297	1	0,040060591	1	0,504701605	4	10	1
Taxi	0,000100567	7	0,002927402	7	Error	7	0,007192210	7	0,042952383	7	Error	7	42	7

## 7. USING GARCH FORECASTS TO PRICE OPTIONS WITH THE BLACK-SCHOLES MODEL

### 7.1 Black-Scholes model: Introduction and intuitive description

The Black-Scholes option pricing formula is the most well-known options pricing formula, and has earned its creators the *Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel*. It is a very elegant formula that produces a price from a potentially complex process. The formula is as follows:

$$C(S, t) = N(d_1)S - N(d_2)Ke^{-r(T-t)} \quad (36)$$

$$d_1 = \frac{\ln\left(\frac{S}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}} \quad (37)$$

$$d_2 = \frac{\ln\left(\frac{S}{K}\right) + \left(r - \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}} \quad (38)$$

$$P(S, t) = Ke^{-r(T-t)} - S + C(S, t) \quad (39)$$

(Black and Scholes, 1973)

$C(S,t)$  is the price of a call option where the underlying spot price is  $S$  and the current period is  $t$ .  $P(S,t)$  is the price of the put option. The other parameters are the strike price  $K$ , the risk-free rate  $r$ , the date of maturity  $T$ , and what this thesis is primarily concerned with  $\sigma$ , the volatility. The reader is asked to use one of the many available resources in order to understand how to use the formula<sup>32</sup>.

Figure 16 crudely acts as a tool for explaining just how the Black-Scholes formula works. As easy as it is to accept the formula and input values, it is probably better to also know why one gets certain prices. In simple terms: the price of the underlying asset follows a Brownian motion, described by the Wiener process – this is the continuous version of the discrete random walk (Taylor, 2007: 354). For the purposes of this example a random walk is used, but by decreasing the time intervals the random walk approaches a Brownian motion (Taylor, 2007: 354). At  $t_0$  the asset is at a certain price, but it faces a continuous probability distribution with regards to where the price will be in the following period. Our knowledge is limited to what is known at  $t_0$ , hence the price could be anywhere along the distribution in  $t_1$ , some places (prices) are more likely than others. For each of these places in  $t_1$ , another distribution can be applied for  $t_2$ . Most places on each of the distributions end up overlapping with many other distributions, but the further from the centre the overlapping probabilities equate to less. For the purposes of the diagram, it can be assumed that the bell curves all encompass a 99% probability. As time goes on, given our information is limited to  $t_0$ , the area we are 99% sure the price will land up grows, with the most likely scenario (given our information) being in the middle. Now, using the example of a call option with a strike price of \$1340, expiring at  $t_4$ , a value needs to be given for the option to buy this asset at this price at that time. Since this option is only worth something if the price actually ends up being greater than \$1340 at  $t_4$ , one is purely interested in the expectation of all prices above this level. To obtain this expectation the resultant distribution is applied in an integral to all the values above \$1340 minus \$1340 and a number is given<sup>33</sup>. This number is the price of the option because risk-neutrality is assumed. Risk-neutrality merely means that an agent always prefers to maximise value, and is not biased towards or against safer bets with disproportionately lower

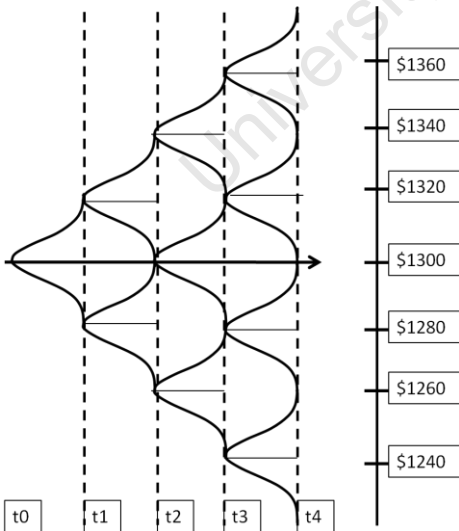
<sup>32</sup> See Black and Scholes (1973) for the original, or visit <http://en.wikipedia.org/wiki/Black%E2%80%93Scholes> for a quick and easy overview.

<sup>33</sup> If one has the probability density function at  $t_4$ , it is a matter of integrating the PDF multiplied by  $x$  (the price) with the integrals limits set from \$1340 to infinity.

payoffs. The same applies to a put option, except the expectation is taken over the other side (below) of \$1340<sup>34</sup>. Clearly a put option (with \$1340 strike) will be a lot more expensive in this case because the expectation of its realised underlying value (\$1340 minus any price below \$1340) is a lot higher.

The Black-Scholes model also takes into account the risk-free rate which values the return of exercising the option in terms of its excess return over this rate. One can now see why an out of the money option is worth more when there is a higher volatility – the expectation of certain prices increases<sup>35</sup>. Also, at the money options are worth less the longer the time to maturity as the expectation of the spot remaining stable lessens. The Black-Scholes formula very elegantly models this theory and allows one to enter but a handful of variables to achieve a precise number. However, understanding this process also exposes the fallibility of the model. It relies on the assumption of a normal distribution, and a constant rate of volatility (and risk-free rate). If volatility is better described by fatter-tailed distributions then this is not accounted for by the model. And, if one expects volatility to gradually increase or decrease then this is also not accounted for – however, it would appear that the slight forecasted changes in the standard deviation is not nearly as big an issue as the assumption of normality. Options should be priced higher because larger deviations should carry higher probabilities of occurrence. This is evidenced by the results of this thesis’ experiment which shows that options with highly deviant strike prices tend to be sold for too low a price, and can lose money for the issuer.

Figure 16: Visual Representation of Black-Scholes formula’s assumption of Brownian motion



<sup>34</sup> Same as above comment, except integral limits are from zero to \$1340.

<sup>35</sup> Higher probabilities are assigned to more extreme prices.

### 7.2 Methodology: GARCH models in Black-Scholes pricing models

In order to test the performance of the different GARCH models in a more practical way, the out-of-sample predictions were applied to option pricing because it explicitly takes volatility into account. A uni-directional strategy is employed where one buys one put and one call option (European). A call is the right to buy an asset at the strike price, and a put the right to sell an asset at the strike price (CFA Institute, 2010: 83). Hence, when the asset ends up above the strike price of the call option, the call option is worth the spot price (on the day of expiry) less the strike price. And, when the asset price ends up below the strike price of the put option, the put option is worth the strike price less the asset price. However, if the asset price ends up being above the put strike, but below the call strike, then neither option is worth anything as one would make a loss if they exercised them. This strategy therefore bets that deviations of a certain magnitude will take place, either positive or negative. This strategy is known as a “long straddle” (The Options Guide, 2011). It is represented in figure 17. The strategy aims to have limited risk (equivalent to the cost of the options) but unlimited reward (upside potential). As can be seen from the figure, the realised maturity price needs to be substantially different from the spot price at  $t_0$ , and even from either strike price in order to ensure that a profit will be made – hence the strategy bets that large deviations will occur. It is implemented for options with one day to maturity using the Black-Scholes formula, for 252 days for each of the GARCH models, for four different bands of deviation from the previous day’s price (1%, 1.5%, 2.5%, and 3.5%)<sup>36</sup>. The total profit over the course of the year indicates the differences between the models. The option buyer’s profit is the option issuer’s loss, and vice versa. In fact, this strategy can also be viewed as a “short straddle” from the issuer’s perspective.

Figure 17: Long Straddle option strategy

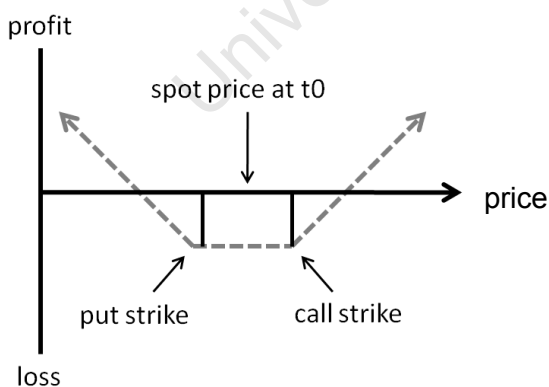
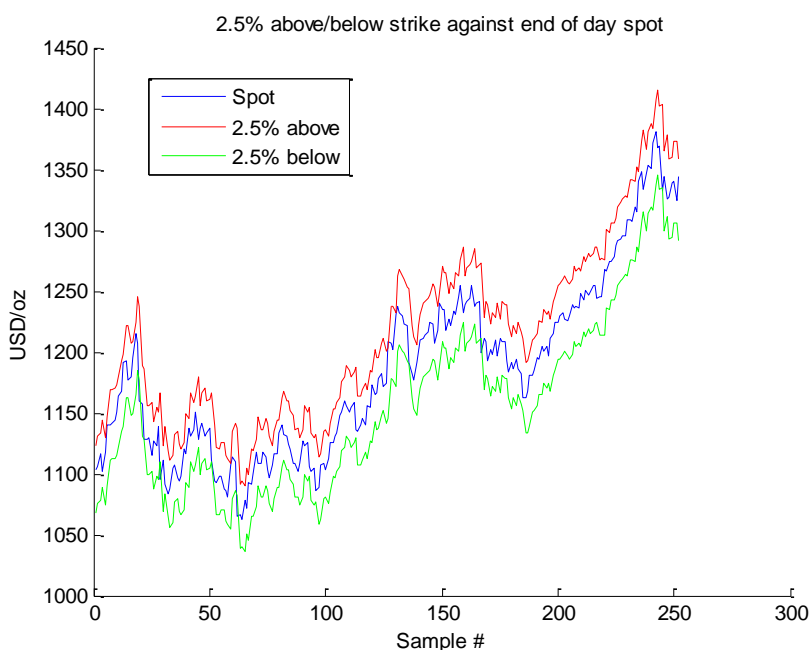


Figure 18 shows the long straddle in action for the sample of 252 observations that are used. It is merely a band of 2.5% above and below the day before maturity’s

<sup>36</sup> This is done by using a loop in the MATLAB package, and the function `blsprice.m`.

spot price, against the day of maturity spot price. When the blue line is either below the green line, or above the red line, either the put or call option is worth something (not necessarily more than the original price of the option). Each strategy for each model therefore is the sum of the cost of both the put and call options for all periods, plus the price of either the put or call option on the day of maturity (if it is more than zero). For each day the one step-ahead volatility forecast is inputted into the Black-Scholes formula. The formula needs annualised volatility, and so the forecasts are multiplied by  $\sqrt{252}$ <sup>37</sup>. The US Treasury bill daily rate is used for each corresponding period.

Figure 18: 2.5% Long Straddle for sample



38

### 7.3 Results of the options strategy experiment

Table 13 shows the profits that the option buyer would have made, had the issuer used the predictions offered by each of the GARCH models. Hence, when the buyer makes a profit, the issuer makes a loss. But, the issuer is the one who chooses the prediction model, and so he should choose that model that maximises his profits, or minimises his loss. Therefore, each model is ranked in terms of the

<sup>37</sup> Annualised standard deviation is the square root of annualised variance, which is  $\sigma\sqrt{252} = \sqrt{\sigma^2 252}$ .

<sup>38</sup> The blue (middle) line is  $P_{t+1}$ , whilst the other two lines are  $1.025xP_t$  and  $0.975xP_t$ . This particular sample begins on the 6<sup>th</sup> of November 2009 and ends 28<sup>th</sup> of October 2010.

issuer’s profit/loss – the model that yields the highest profit for the buyer is considered the least effective and is ranked lowest.

The interesting thing about these results is how much they conflict with the results of the statistical loss functions. Whereas the models that took account for leverage obtained the best fit for all horizons when tested with the statistical loss functions, the simpler GARCH(1,1) models performed the worst from the purchasers perspective (or alternatively, they would have provided the issuer with the highest profit margin).

Figure 19: Forecast comparison for sample

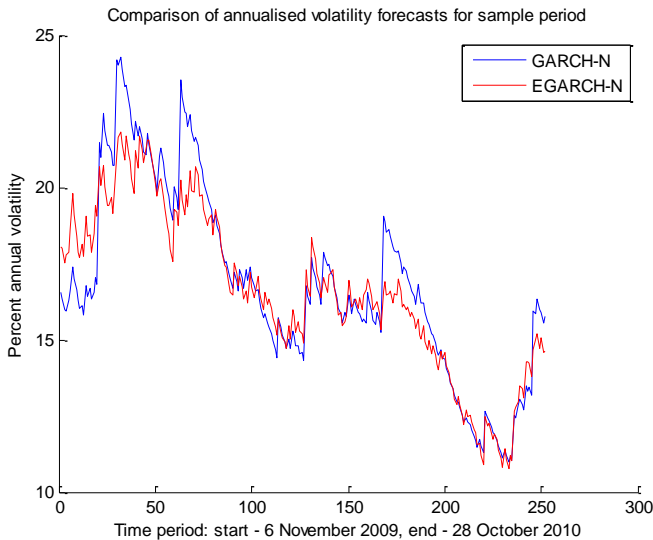


Figure 19 illustrates the difference between the step-ahead forecasts of a GARCH(1,1) model with an EGARCH model, both using the normal distribution. The EGARCH appears to be more conservative, with lower peaks, but often produces higher forecasts. However, the forecasts of the EGARCH are, on average, suggesting lower option prices because of the lower volatility forecasts in certain areas<sup>39</sup>. Figures 20 to 23 illustrate the shape of the profits made by the long straddle strategy over the 252 periods. As can be seen, in all the figures the GARCH model (blue line) is usually at the bottom, which corresponds to charging the buyer higher prices on average for the options. Although this test is highly theoretical, because options prices are ultimately decided by the market, it does serve to recommend to an issuer that he may do well to consider using the GARCH(1,1) in preference to

<sup>39</sup> This is apparent when one looks at Vega, the partial first derivative of the option prices with respect to volatility. This value is positive for out of the money options.

$$\frac{\delta C}{\delta \sigma} = SN'(d_1)\sqrt{T-t}$$

the EGARCH or GJR-GARCH as the basis for pricing OTC options, purely because it appears (from this sample) that it would work out to be more profitable over the long run by balancing out the occasional extreme losses of his short straddle<sup>40</sup>.

Table 13: Results of long straddle strategy

	Profit								Total	Rank
	1%	Rank	1,50%	Rank	2,50%	Rank	3,50%	Rank		
GARCH-T	-\$37,45	6	\$20,41	6	\$34,84	5	\$10,56	5	22	5
GARCH-N	-\$32,89	5	\$21,71	5	\$34,36	6	\$10,38	6	22	5
EGARCH-T	-\$20,05	4	\$34,49	4	\$40,37	3	\$11,70	1	12	3
EGARCH-N	-\$7,17	1	\$40,34	1	\$40,80	1	\$11,66	3	6	1
GJR-T	-\$14,71	3	\$36,89	2	\$40,50	2	\$11,69	2	9	2
GJR-N	-\$12,40	2	\$36,73	3	\$39,68	4	\$11,49	4	13	4

	Option Cost								Total	Rank
	1%	Rank	1,50%	Rank	2,50%	Rank	3,50%	Rank		
GARCH-T	\$602,44	6	\$249,08	6	\$32,49	5	\$3,18	5	22	5
GARCH-N	\$597,89	5	\$247,78	5	\$32,98	6	\$3,37	6	22	5
EGARCH-T	\$585,04	4	\$235,00	4	\$26,97	3	\$2,04	1	12	3
EGARCH-N	\$572,16	1	\$229,16	1	\$26,54	1	\$2,09	3	6	1
GJR-T	\$579,70	3	\$232,60	2	\$26,84	2	\$2,06	2	9	2
GJR-N	\$577,39	2	\$232,77	3	\$27,65	4	\$2,26	4	13	4

In figures 20 to 23, the differences that the various strike price bands make to the profit/loss are clearly apparent. When the strike band is low (1% which is less than the standard deviation), the frequency of profiting from the strategy is more, but the cumulative cost of the options is also a lot more – the precise figures can be seen in tables 13. However, at the other extreme, when the strike band is quite high the frequency is lower and the profit per frequency is also lower, but the cumulative options cost is much lower. This means that there is an optimum strategy for a trader utilizing the long straddle – in this case it is the 2.5% band as it yields the highest profit across all forecasts. This is also the worst band for the short straddle strategy.

<sup>40</sup> Nick Leeson famously bankrupted Barings Bank in 1995 after using the short straddle strategy on Japanese stocks (Time, 1995).

Figure 20: 1% Straddle

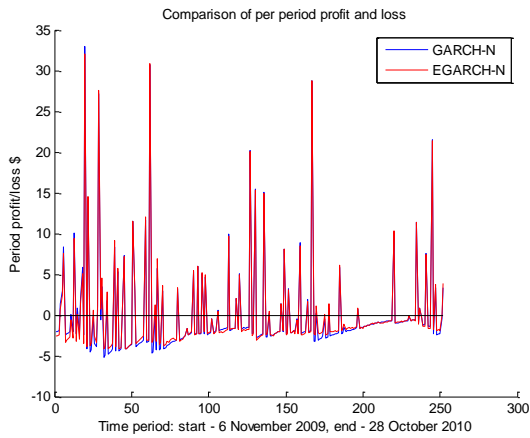


Figure 21: 1.5% Straddle

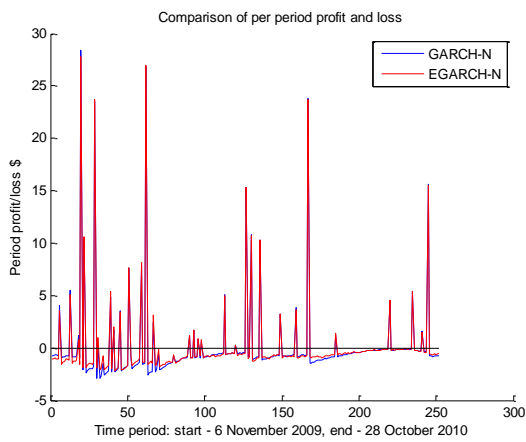


Figure 22: 2.5% Straddle

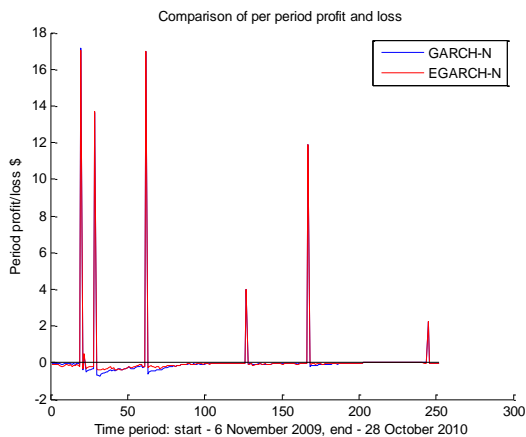
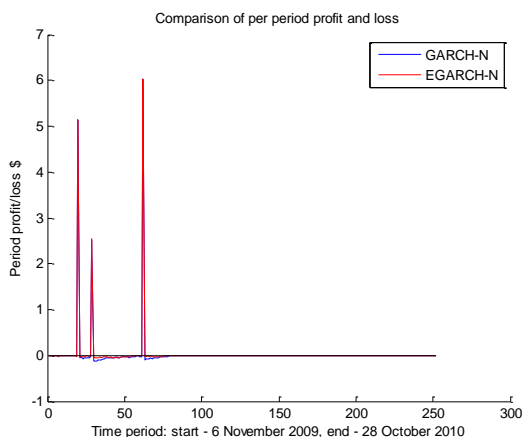


Figure 23: 3.5% Straddle



#### 7.4 Some caveats

The above methodology is entirely theoretical. In reality it would be very difficult to get someone to sell you an option as small as some of the prices from the 3.5% straddle. In addition, although traders make use of the Black-Scholes model, they will have different methods of correcting for the model's shortcomings – prices in the real market will differ in many ways to this thesis' theoretical prices. There are also other shortcomings of the Black-Scholes model that are widely recognised. These include the assumption of constant interest rate until maturity, no transaction costs, ability to short sell, no taxes, continuous trading of the underlying and option, and no arbitrage (Taylor, 2007: 373). These assumptions may or may not be true in any one market, and if they are not true there are ways of dealing with them such as specially weighted interest rates (Hull, 2009). However, dealing with these shortcomings are beyond the scope of this paper, but it is recommended that future research investigate what effects these have in the gold market and how they relate to the use of GARCH models in asset pricing.

An important thing to note is that the model used in this thesis' estimation may have been misspecified. Biger and Hull (1983: 25) state that the traditional Black-Scholes formula needs to be adjusted for currencies, where the underlying asset has a continuous payout – the interest rate. Although gold isn't specifically a currency, it is often treated as such, and does indeed have an interest rate – the gold lease rate. This rate is the rate at which a financial institution is prepared to lend gold at (Whaley, 2006). It is, however, not quoted in the market as it is usually available to larger market participant only. The implied rate can be found at [www.lbma.org.uk](http://www.lbma.org.uk) (The London Bullion Market Association) – it is calculated as LIBOR (London Inter Bank Offered Rate) minus GOFO (Gold Forward Offered Rate). Unfortunately, there were complications with this rate for the particular out-of-sample used in the estimations – the implied rate was negative, which is a mathematical anomaly. For simplicity's sake the usual Black-Scholes equation was

stuck to. However, if further research were to be done with gold data, it is worth noting that the revised model is as follows:

$$C(S, t) = N(d_1)Se^{-rf(T-t)} - N(d_2)Ke^{-rd(T-t)} \quad (40)$$

$$d_1 = \frac{\ln\left(\frac{S}{K}\right) + \left(rd - rf + \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}} \quad (41)$$

$$d_2 = \frac{\ln\left(\frac{S}{K}\right) + \left(rd - rf - \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}} \quad (42)$$

$$P(S, t) = Ke^{-rd(T-t)} - S + C(S, t) \quad (43)$$

(Biger and Hull, 1983: 25)

The model is very much the same as the traditional model, except for the continuous payout rate  $rf$  which has been placed within the equations. This variable represents the foreign interest rate, if one were to be buying an option where the foreign currency was the underlying. In this case, gold is the underlying currency. The domestic rate of interest for the currency that the option is priced in (US Dollars in this case) is just the variable  $rd$ . It is not expected that the results would change drastically with this thesis' estimation purposes as the options are so short dated that all interest rate effects are marginal.

## 8. CONCLUSION

Three different GARCH-type models, with both normal and student's  $t$  distributions, were tested for their ability to forecast the volatility of the gold price. The models were tested in both a loss function framework, and an option pricing framework.

It was found that the models that took into account the differing effects of positive and negative return data, demonstrated better forecasting ability when statistical loss functions were used to measure fit.

However, when an options straddle strategy was tested with the one-step ahead forecasts, it was found that the regular GARCH models provided better results, in the form of lower losses and higher profits, for the issuer of the options.

The results are therefore mixed, and possibly inconclusive. The out-of-sample time period was for a full working year which was intended to get a more comprehensive perspective on the results.

There are problems with both the statistical loss function measurements, and the option pricing measurements. The loss functions do not tell us the monetary implications of the forecasts, whilst the option methodology had many noted flaws in assumptions and relation to reality.

In light of these issues, it is concluded that the observations from the statistical loss functions were more relevant as they were more comprehensive in scope having used four different forecast horizons, whereas the options results used just the one day horizon. Therefore, it can be said that gold exhibits volatility clustering, and using GARCH models can indeed be useful in forecasting its volatility. In addition, the volatility in gold reacted to negative returns more so than positive returns, evidenced in the better performance of EGARCH and GJR-GARCH over the standard GARCH. The fatter tailed student's  $t$  distribution also outperformed the standard normal distribution, and the best performing model in all time horizons was the GJR-GARCH with a student's  $t$  distribution.

Further research is recommended to build on these findings, to test more intricate versions of the models and to get robust results in pecuniary terms.

University of Cape Town

## REFERENCES

- BIGER, N. And HULL, J. 1983. The Valuation of Currency Options. *Financial Management*. 12, 1: 24 – 28.
- BLACK, F., 1976. "Studies of Stock Market Volatility Changes," 1976 Proceedings of the American Statistical Association, Business and Economic Statistics Section: 177-181.
- BLACK, F. and SCHOLES, M., 1973. The Pricing of Options and Corporate Liabilities. *The Journal of Political Economy*. 81, 3: 637 – 654.
- BOLLERSLEV, T., 1986. Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*. 31: 307 – 327.
- BOLLESLEV, T. and ANDESON, T., 1998. Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economics Review*. 39, 4: 885 – 905.
- BOLLERSLEV, T., CHOU, R. and KRONER, K., 1992. ARCH modelling in finance: A review of the theory and empirical evidence. *Journal of Econometrics*. 52: 5 – 59.
- BOLLERSLEV, T., ENGLE, R. and NELSON, D., 1994. "ARCH models" in R.F. Engle and D. McFadden, eds., *Handbook of Econometrics Volume IV*. Amsterdam: North Holland Press.
- CHRISTOFFERSEN, P. and JACOBS, K., 2004. Which GARCH model for option valuation? *Management Science*. 50, 9: 1204 – 1221.
- CFA INSTITUTE, 2010. *Derivatives and Alternative Investments*. Cape Town: Pearson.
- ENGLE, R.F., 1982. Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*. 50, 4: 987 – 1007.
- ERSTE GROUP, 2010. *Special Report Gold*. Erste Group. [online]. Available: [www.usagold.com/publications/erste-2010.pdf](http://www.usagold.com/publications/erste-2010.pdf) [Accessed 20 January 2011].
- FUNG, W.K.H. and HSIEH, D.A., 1991. *Empirical Analysis of Implied Volatility: Stocks, Bonds and Currencies*. Duke University: Department of Finance, Fuqua School of Business.
- GHYSELS, E., HARVEY, A. and RENAULT, E., 1996. "Stochastic Volatility" in G.S. Maddala and C.R. Rao, eds., *Handbook of Statistics Vol. 14, Statistical Methods in Finance*. Amsterdam: North Holland Press.
- GLOSTEN, L.R., JAGANNATHAN, R. and RUNKLE, D.E., 1993. On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *The Journal of Finance*. 48, 5: 1779 – 1801.
- HANSEN, P.R. and LUNDE, A., 2005. A forecast comparison of volatility models: Does anything beat a GARCH(1,1)? *Journal of Applied Econometrics*. 20: 873 – 889.
- HOSKING, J.R.M. and WALLIS, J.R., 1987. Parameter and Quantile Estimation for the Generalised Pareto Distribution. *Technometrics*. 29, 3: 339 -349.

- HULL, J. C., 2009. *Options, Futures, and Other Derivatives*. Englewood Cliffs, NJ: Prentice-Hall.
- HSIEH, D.A., 1991. Chaos and Nonlinear Dynamics: Application to Financial Markets. *Journal of Finance*. 46: 1839 – 1877.
- MANDELROT, B., 1963. The Variation of Certain Speculative Prices. *Journal of Business*. 36: 394-419.
- MARCUCCI, J., 2005. Forecasting Stock Market Volatility with Regime-Switching GARCH Models. *Studies in Nonlinear Dynamics & Econometrics*. 9, 4: article 6.
- NELSON, D.B., 1991. Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica*. 59, 2: 347 – 370.
- POPPER, K.R., 1972. *Objective Knowledge: An Evolutionary Approach*. Oxford: Clarendon Press.
- SHEPARD, N., 1996. "Statistical Aspects of ARCH and Stochastic Volatility" in D.R. Cox, D.V. Hinkley and O.E. Barndorff-Nielson, eds., *Likelihood, Time Series with Econometric and Other Applications*. London: Chapman and Hall.
- TALEB, N.N., 2010. *The Black Swan: The Impact of the Highly Improbable (2e)*. Rosebank, Johannesburg: Penguin Books.
- TAYLOR, S.J., 2007. *Asset Price Dynamics, Volatility and Prediction*. Princeton, New Jersey: Princeton University Press.
- THE OPTIONS GUIDE, 2011. *Option Straddle (Long Straddle)*. The Options Guide. [online]. Available: <http://www.theoptionsguide.com/long-straddle.aspx> [Accessed 12 January 2011].
- TIME, 2005. *Nicholas Leeson: GOING FOR BROKE*. Time Magazine. [online]. Available: <http://www.time.com/time/magazine/article/0,9171,982669-1,00.html> [Accessed 13 January 2011].
- WHALEY, R.E., 2006. *Derivatives: Markets, Valuation and Risk Management*. New York: Wiley.

## ACKNOWLEDGEMENTS

I would like to thank Kevin Kotze for his role as supervisor on this thesis – his expertise and advice were invaluable. In addition, Peter Searle and Robert Price were group members with me on a term paper that was the initial inspiration for this work – thank you to them.