NL Samouilhan and G Shannon*

Forecasting volatility on the JSE

1. INTRODUCTION

Economics, as a school of thought, is chiefly concerned with the optimal allocation of scarce resources in an attempt to maximise (or minimise) some function. In the large sub-school of financial economics, the scarce resource is capital, the goal is to maximise the asset base and the optimal allocation happens chiefly through the financial markets. The correct discriminator for the allocation of capital, given the maximisation function, is the rate of growth (or return) it offers. However, as the allocations inevitably occur in an environment pervaded with uncertainty, an adjustment must be made for the risks associated with every return. The resultant risk-return relationship is at the heart of financial theory and practice, with higher, more certain returns always being preferred to lower, less certain returns. Financial economics then, as a sub-school of economics, reduces to the study of the risks and returns of the competing uses of capital.

In the equity market, a major capital allocation mechanism, the return is defined as the movement in an asset's price, and the risk as the variance of those returns: a perspective stretching back to Markowitz (1952) and Tobin (1958). In practice, as the allocations of capital (investment decisions) are invariable forward looking, it is the expected risks and expected returns which must be studied, and for this accurate forecasts of both are required.

Of central importance to the ability to forecast any series is the degree of persistence in it, and in this regard the risks and rewards of equity returns differ drastically. For equity returns, even weak market efficiency precludes inertia in asset prices in equilibrium; arbitrage conditions ensure that it should be impossible to forecast the future behaviour of an equity asset based on the analysis of its past behaviour. This, the very weakest assumption of market efficiency, renders forecasts of returns based on its past behaviour impossible[1]. In sharp contrast to this there is now wide-spread evidence that equity risks, defined in the literature from Markowitz (1952) onwards as the variance of the equity returns, have considerable inertia. This persistence implies that the lagged values of domestic equity volatility will provide significant information regarding the future level of volatility, and hence forecasts are indeed possible. The natural question then, and the one that this paper attempts to answer, is which specification of volatility provides a better guide to future equity volatility.

Specifically, the aim of this paper is to investigate the comparative ability of three different types of volatility forecasts. Firstly, the ability of Autoregressive Conditional Heteroskedasticity (ARCH) process models to provide good estimates of equity volatility is well documented, and forecasts based on these processes often provide good estimates of future volatility (Blair, Poon and Taylor 2001). Secondly, using iterative procedures it is possible to impute the volatility priced into a derivative contract. This implied volatility is often taken to be the 'market' forecast of volatility. Using the new SAVI[2] (Safex inter-Bank Volatility Index) series of implied volatility on the TOP40 Index developed by Cadiz Securities, this 'market' forecast can be rigorously evaluated for the first time. Lastly, forecasts of volatility based purely on historical volatility, such as a random walk (naive) or moving average forecast, are still popular and were the dominant form of volatility forecasts before ARCH process models. Looking at the forecasting ability of these relatively simple models, in both a relative and absolute manner, will provide much insight.

This paper is laid out as follows. Section 2 discusses the equity volatility forecasting literature in order to put this domestic study into perspective. Section 3 then continues by explaining what is meant by 'equity volatility' and how this is defined in practice, while Section 4 defines and explains the various volatility models that are evaluated. Section 5 presents the econometric methodology and data used in this study. Sections 6 and 7 present the key findings: Section 6 discusses the in-sample fit of these volatility

*Respectively, School of Economics, UCT and Equity Derivative Research, Cadiz FSG
Email: nicholas.samouilhan@uct.ac.za; grant.shannon@cadiz.co.za

The authors are indebted to Dr CP Van Walbeek, Professor J Fedderke, Assoc. Professor M Wittenberg, participants in the UCT School of Economics seminar series and two anonymous referees from the journal for many helpful comments made on an earlier draft, and to Mr M. Pillay for help with the gathering and construction of the data. Any errors or omissions however remain the authors' responsibility.

[1] It should be acknowledged that there is some debate regarding whether equity markets are indeed weak-market efficient, see for e.g. the domestic research on domestic equity overreaction by Cubbin, Eidne, Firer and Gilbert (2006). This debate is however not materially important here. The existence of weak market efficiency is not central to this paper, and is discussed here only to highlight the difference in inertial forecastabililty between variances and returns. The presence of weak market inefficiency would change this discussion from one of absolutes to one of relative degrees, but even by the weak-market inefficiency literature the difference is sufficiently large that the point of radically different inertial forecastabililty remains largely the same.

[2] This is the approximate domestic equivalent of the CBOE's VIX, as discussed further on in the paper in more detail.

specifications, while Section 7 discusses their out of sample forecasting ability. Section 8 concludes.

2. LITERATURE REVIEW

The research into the forecasting ability of different ARCH process models is extensive, though it does tend to be overwhelmingly focused on the equity markets of a few developed countries. While the conclusions[3] from this literature are mixed, a few salient points can still be distilled from the studies on index volatility. Firstly, it appears that equity volatility can to some degree be forecasted. This is in sharp contrast to the forecasting literature regarding the equity return, where this aspect is fiercely contested. Secondly, regarding the ability of the different ARCH process models to explain volatility: Generalised ARCH (GARCH) models tend to provide both superior in- and out-of-sample fit than ARCH models, while accounting for asymmetry tends to be desirable at times.

With regards to implied volatility, all but one study (Canina and Figlewski, 1993) find that it provides useful information about future volatility. Turning to the comparative forecast records approximately half (out of 39) of the ARCH process versus Implied Volatility papers surveyed find that ARCH is better, with the other half finding that the converse. Historical models are better than Implied Volatility in 24% of the 34 papers surveyed, with 76% finding that Implied beats Historical. Lastly, ARCH processes outperform Implied in only one paper, with 17 others finding the opposite is true.

The justification for this study is three-fold. Firstly, given the importance of volatility to domestic financial markets this study hopes to provide practical information as regards to the specification of the domestic equity volatility. Secondly, the comparison of the new SAVI with the ARCH process and historical volatility models will provide information on whether the use of these models could improve the volatility forecasting ability of the domestic market. Lastly, as Bollerslev, Chou and Kroner (1992) and Engle (2001) bemoan, with few exceptions almost all of the research into volatility has used data from the US, the UK and the Japanese markets. The study of additional, novel equity markets such as South Africa will allow for a greater understanding of the volatility forecasting process itself as country specific factors are diversified out. The JSE is uniquely positioned to provide such an understanding: it is both relatively large, liquid and open, yet has a composition that is different to that of the developed market bourses of New York, London and Tokyo.

[3] This summary is based, in part, on the survey of the literature of Poon (2005).

3. EQUITY VOLATILITY AS A LATENT VARIABLE

Before the different volatility models are discussed, it must be made clear what they are attempting to forecast, and hence what they must be evaluated against. Standard financial theory holds that an equity's price reflects the net present value of the expected dividends associated with such an equity. As such, the share price will only change as new, relevant information is priced in. Assuming only weak market efficiency and that the news arrival process is random, equity prices will then behave according to a stochastic Markov process. The most widely used continuous-time, continues-variable specification of this is given by the sub-martingale series:

$$\delta S = \mu dt + \sigma dz \qquad \dots (1)$$

Where δS is the change in the price (S) of the equity (actual return), μ is the expected return of the asset over time and σ is the variance of the equity. The variance of the returns is generated by the specification σdz, where dz is a Weiner process with a standardised normal distribution, a mean of 0 and a standard deviation of 1. Given the above specification, the distribution of the returns on the asset is given by:

$$\delta S \sim (\mu, \sigma)$$

This paper is concerned with the forecasting of σ, the volatility (risk) of the equity returns. Equity volatility is not directly observable, and (as a consequence) is an inherently latent measure. In practice, discreet price changes must be used as a proxy for such volatility, and, for the proper evaluation of a volatility forecast, it is imperative that this is a reasonable proxy. For this, the equity volatility literature has traditionally employed the method of using squared daily returns as the proxy. While not ideal due to the idiosyncratic nature of daily returns, this volatility proxy was also employed for three reasons. First, as the usual measure of equity volatility its use here will allow greater comparability of the results. Secondly, the ARCH process models must be estimated using daily returns, and the use of these same returns to generate the volatility greatly simplifies the analysis and interpretation. Lastly, most other measures of volatility may not be used given that the volatility estimated from the ARCH series is defined as movement around the mean, whereas these measures (realised volatility and the High-Low method for example), define volatility as the total movement on a certain day. Stated differently, under ARCH models if the equity return is equal to its sample (non-zero) mean, then there is by definition no volatility. With the H-L and realised volatility, this or any movement at all is defined as volatility. However, while this is an important theoretical distinction this is in practice a trivial point: the return on a day is likely to be dominated by the excess return component (see below), rendering the measures approximately equal.

4. VOLATILITY FORECASTS

4.1 ARCH process forecasts

ARCH Process models model volatility as being the residual to some mean estimation function. The fundamental mechanism whereby they work rests on two related properties: that the variance *σ* in equation (1) is far from constant in financial time series, and that there tends to be a large degree of 'memory' or persistence in such variance levels.

In regard to the first property, equation (1) must then restated as:

$$\delta S = \mu dt + \sigma_t dz \qquad \dots (2)$$

With the variance at time t being conditional on time t. From here on this paper will refer to this as the conditional volatility or conditional variance of the series.

The second property implies that this conditional volatility tends to exhibit large degrees of clustering (or persistence) over time. A period of high volatility tends to be followed by another period of high volatility, while low volatility periods are followed by equally low volatility periods. As such, the conditional variance in equation (2) can be modelled as:

$$\sigma_t^2 = f(\sigma_{t-i}^2) \qquad \dots (3)$$

that is, as (partly) a function of its past values.

Returning to the return specification, through the discrete sampling of equation (2) the one-period return for period T can be decomposed into the expected and unexpected part:

$$r_t = \mu + \varepsilon_t \qquad \dots (4)$$

Where μ and ε_t represent, respectively, the expected and unexpected components of the sampled return, and ε_t is the sampled observation of σdz at time t. Note that the expected value of ε_t remains zero due to the *dz* term, implying that the expect return is equal to μ. As volatility is defined as the movement of the returns around the mean, then equation (4) of the conditional variance can be restated as:

$$\varepsilon_t = r_t - \mu \qquad \dots (4')$$

which is the discrete sampled representation of the volatility at time t. Using this definition of excess returns as the proxy for volatility, equation (3) becomes:

$$\sigma_t^2 = f(\varepsilon_{t-i}) \qquad \dots (5)$$

Equation (5) above represents the most general specification of the ARCH processes, which attempt to model the variance of the returns (risk) based on the previous unexpected return. In modelling South African equity volatility, this paper will evaluate three main specifications of $f(\varepsilon_{t-i})$, namely ARCH, GARCH and TARCH. These are discussed further immediately below.

(i) ARCH

The first and simplest ARCH specification is the ARCH model of Engle (1982), where the forecasted conditional variance for period t+1 is modelled as simply being a function of some long term mean and the square of the previous excess or unexpected return:

$$\hat{\sigma}_{t+1}^2 = \omega + \alpha\varepsilon_t^2 \qquad \dots (6)$$

with ω being the time-invariable mean and ε_t^2 being the square of the excess return in equation (5) above.

(ii) GARCH(p,q)

The standard ARCH model above can be generalised as by Bollerslev (1986) by including the conditional variance(s) of the previous period in addition to the time-invariable mean and the excess return from the previous period(s). This model, known as a GARCH(*p,q*) model, contains GARCH terms to the lag of *p* and ARCH terms to the lag of *q*:

$$\hat{\sigma}_{t+1}^2 = \omega + \sum_{i=0}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=0}^{p} \beta_i \sigma_{t-i}^2 \qquad \dots (7)$$

While the specification of the GARCH(*p,q*) model allows for an infinite number of ARCH and GARCH lags, in practice most studies adopt low orders or 2 or less for both *p* and *q* (Bollerslev, Chou and Kroner, 1992). In this study of GARCH forecasting ability, specifications of (1,1), (2,1), (1,2), and (2,2) will be investigated.

(iii) Threshold ARCH (TARCH)

A regular empirical finding in the ARCH equity literature is of an asymmetrical response of volatility to past returns. Negative excess returns (the unexpected return ε_t in equation (4)) are found to be followed by significantly higher volatility than positive returns of the same magnitude. The standard ARCH and GARCH specifications, in ignoring the direction of past errors, do not account for this possibly useful information. In contrast, the Threshold ARCH (TARCH) model of Glosten, Jaganathan and Runkle (1993) includes a dummy variable, d_t, to distinguish negative returns from positive returns. With the inclusion of such an asymmetric effect into a standard GARCH (*p,q*) model, the TARCH(*p,q*) forecast for period t+1 is specified as:

$$\hat{\sigma}_{t+1}^2 = \omega + \sum_{i=0}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=0}^{p} \beta_i \sigma_{t-i}^2 + \gamma d_t \varepsilon_t \qquad \dots (8)$$

where

d_t is 1 if $\varepsilon_t < 0$, and 0 otherwise.

The value γ is therefore the extent that negative returns affect future conditional volatility by more than a positive return of the same magnitude. This paper will investigate both a TARCH(1,1) model and a TARCH(2,2), which are the corresponding asymmetric versions of the GARCH(1,1) and GARCH(2,2) models, respectively.

4.2 Historical volatility forecasts

While the ARCH process models construct estimates of volatility using the information provided by the mean generating process, historical volatility uses the information from volatility measures directly. As such, they are far easier to construct and manipulate, which partly explains their continued use and attraction. This paper will look at two versions of historical volatility: the naïve (random walk) forecast and the 5 day moving average. The naïve forecasts has the volatility for period (t) being the volatility that existed at period (t-1). As such, the forecasted volatility is simply the 1 period lagged volatility, in this case the squared excess returns. The 5 day moving average forecast, in contrast, weights the volatility (again, proxied by the squared excess returns) of the previous 5 days equally. Clearly, these forecasts should be secondary in forecasting power to the ARCH models, which include both a measure of mean reversion and some approximation of the volatility term structure. However, given their historical importance and so as to act as a reference base, they are still generated.

4.3 Implied volatility forecasts

The models discussed thus far depend on historical return data to generate volatility forecasts. There is however an alternative approach to volatility forecasting that relies on implied volatility data obtained from option prices. This implied volatility is a measure of future asset price return uncertainty over a specific time period, as priced in by the market. As such, it is usually defined as the forward looking market forecast of volatility.

An implied volatility index is a transparent representation of such a market forecast. The most referenced implied volatility forecast is the Chicago Board Options Exchange Volatility Index, usually referred and known by its ticker symbol 'VIX'. This volatility forecast is a daily measure of the expected volatility over the next 30 days that has priced into the S&P 500 index options. The current VIX uses a strike based weighted average of option prices to calculate the implied volatility index.

The South African Futures Exchange (SAFEX), in collaboration with Cadiz Securities, has built a South African equivalent implied volatility index called the SAVI – The South African Volatility Index.
The SAVI represents a daily measure of TOP40 at-the-money implied volatility over the following three months. Its calculation is such that, when compared with the current VIX formulation, it represents a lower bound for three month volatility as a whole. For further discussion regarding the construction of the SAVI, see Cadiz (2006).

The SAVI was constructed using the original methodology proposed by Whaley (2000). This approach is simpler than the reconstructed CBOE (2003) formula. Adjustments are made, however, to Whaley's method to accommodate the less liquid TOP40 option market i.e. only options within a neighbourhood of "at the money" are considered in the index construction process.

On any given day the SAVI (as quoted by SAFEX) is an annualised measure of three month implied volatility. For this study, the SAVI is de-annualised by converting it into a daily volatility measure. This is done through the standard practice of dividing the annualised SAVI measure by the square root of 252, the number of trading days in a year. This daily volatility reading then generates forecasts according to the specification:

$$\hat{\sigma}_{t+1}^2 = SAVI_t^2 \qquad \dots (9)$$

5. DATA AND METHODOLOGY EMPLOYED

5.1 Data

The data used in this study was drawn from the McGregor database and SAFEX/Cadiz.

The SAFEX/Cadiz data set comprises daily readings for the SAVI index. This is an implied volatility measure and is calculated from near and next-near SAFEX option data. Only at-the-money TOP40 options that display sufficient liquidity are used in the historic index construction process. More detail is provided in Cadiz (2006).

For the purposed of this research at-the-money represents options that are close to at the money for the expiry concerned. This includes any call (put) option with a delta between 45 and 55 (-45 and -55).

The McGregor data set consists of the daily closing levels of the TOP40 index.

The sample period under investigation runs from 01/02/2004 to 28/09/2006, which translates into 682 actual trading days in total. The dearth of reliable information needed to impute the implied volatility necessitated this relatively short sample period. Ensuring stationarity, the daily closing levels were transformed into daily returns, with these daily returns being defined in the standard way as the natural logarithm of the ratio of the consecutive daily closing levels.

5.2 ARCH process econometric formulation

Operationally, equation (4) was estimated for the ARCH process models using Maximum Likelihood according to the specification:

$$r_t = \alpha + \varepsilon_T \qquad \ldots (10)$$

Note that here, significantly, the value of *α, (*the expected return), is imposed. This was necessary as the maximum likelihood estimation of the above equation will yield different values of *α* depending on the specification of the ARCH process term. As volatility is defined here as the movement of the returns around the mean, the value for *α* was defined as being the daily mean return over the sample period[4]. For the TOP40 over the period under review, this is 0,0111%. Only in this way will an evaluation of the various ARCH process forecasts be comparing like with like.

The error, ε_t (the excess return amount (r_t - α)) has the distribution:

$$\varepsilon_t \sim (0, \hat{\sigma}_t^2) \qquad \ldots (10')$$

and

$\hat{\sigma}_t^2$ being specified according to the respective ARCH model outlined above.

A further final complication in the econometric estimation of is that a salient, well documented property of the distributions of financial time series such as the TOP40 index returns is that of fatter tails than the normal distribution. There are two potential solutions to this problem that are regularly used in the literature. The first involves assuming and using a more likely distribution for the returns, such as the Student-t or Generalised Exponential Distribution (GED). Both of these distributions have been used in the volatility literature; see, amongst others, de Jong, Kemna and Kloeck (1990) for the ARCH use of the Student-t distribution, and Nelson (1989) for the use of GED. The alternative solution, and the one that this paper uses, is to simply assume a Normal Distribution and then employ Bollerslev and Wooldridge (1992) quasi-maximum likelihood (QML) covariances and standard errors to generate the efficient estimators. This procedure of employing robust standard errors is used by, amongst many others, Andersen and Bollerslev (1998) and Blair, Poon and Taylor (2001) in their papers of volatility forecasting, and has the major advantage of being robust against a possible misspecification of the distribution.

6. IN SAMPLE ESTIMATES AND FIT

Table 1 provides the in-sample results of the various ARCH process models. In addition to the significance, coefficients and robust standard errors of the different variable fits, the table also provides the results of the adjusted Coefficient of Determination (R^2adj) measure. Note that this measure is for the fit of the volatility, and not the mean; it is a measure of how well the respective models explain the equity volatility. As such, it applies only to the second, and not the first, moments of the functions.

With the simplest ARCH process specification, the original ARCH model, both the mean reversion and the ARCH effect are significant at the 5% level. However, as can be seen from the adjusted R^2 test this specification does a relatively poor job modelling the volatility, explaining less than 2% of the domestic equity volatility.

The GARCH models in contrast provide a far better fit of the volatility. Note that the additional inclusion of just the one period lagged conditional variance (the GARCH(1,1) model) improves the information content by over 15%. Interestingly, the GARCH(1,2) and GARCH(2,1) models have the approximately the same explanatory power, at 17,8% implying that the one period lagged excess return and one period lagged conditional variance provide approximately the same incremental modelling power. The GARCH(2,2) provides the best in-sample fit of all the symmetric GARCH models, explaining just under 19% of the equity volatility.

[4]Note that this then implies that a correctly specified ARCH process model will provide optimal forecasts of the squared excess daily returns (r_t - α)^2.

Table 1: In sample model parameters for ARCH specifications: 01/02/2004 to 28/09/2006

Model	Parameters						
	ω	β_1	β_2	β_3	β_4	β_5	R^2_{adj} ‡
ARCH	0,0001** (0,0001)	0,1397* (0,0789)	-	-	-	-	0,019
GARCH(1,1)	0,0000** (0,0000)	0,1019** (0,0271)	0,8829** (0,0238)	-	-	-	0,173
GARCH(2,1)	0,0001* (0,0000)	0,1201** (0,0510)	0,6865 (0,6550)	-	0,1777 (0,6006)	-	0,178
GARCH(1,2)	0,0000** (0,0000)	0,0114 (0,0420)	0,8613** (0,0339)	0,0106** (0,0578)	-	-	0,178
GARCH(2,2)	0,0001 (0,0000)	0,0072 (0,0397)	1,1601 (0,5066)	0,0833** (0,0834)	-0,2657 (0,4430)	-	0,188
TARCH(1,1)	0,0001** (0,0000)	0,0311 (0,0396)	0,8905** (0,0251)	-	-	0,1133** (0,0545)	0,179
TARCH(2,2)	0,0000** (0,0000)	-0,0680** (0,0192)	0,6305 (0,4521)	0,1144** (0,0510)	0,2173 (0,4026)	0,1427** (0,0528)	0,172

$r_t = \alpha + \varepsilon_t$

$\varepsilon_t \sim N(0, \hat{\sigma}^2_T)$

$\hat{\sigma}^2_T = \omega + \beta_1\varepsilon^2_{t-1} + \beta_2\sigma^2_{t-1} + \beta_3\varepsilon^2_{t-2} + \beta_4\sigma^2_{t-2} + \beta_5 d_{t-1}\varepsilon^2_{t-1}$

where d_{t-1} is 1 if $\varepsilon_{t-1} < 0$, and 0 otherwise

α is defined as mean log daily return over the sample, 0.00111

Notes: A dash (-) denotes that the parameter was not included in the specific regression estimation
Bollerslev and Wooldridge Robust Standard Errors in Brackets
‡ Note that this refers to the variance equation and not the mean
* *Significant at the 5% level
* Significant at the 10% level

The results from fitting the asymmetric TARCH models show that, in keeping with the international literature, the asymmetrical response of volatility to the direction of the return is found to be significant. The positive coefficients on the negative previous return dummy variable, 0,133 (TARCH(1,1)) and 0,1427 (TARCH(2,2)), indicate that a negative market return on the JSE leads to higher subsequent volatility than a positive return of the same magnitude. However, the adjusted R^2 shows that, when compared to their symmetric versions, the inclusion of the asymmetric term provides no significant incremental information regarding the volatility generating process. In this regard, note that the adjusted R^2 'penalises' these asymmetric models for the additional explanatory terms, and as such it would only increase if the included threshold term significantly increases the model fit. This explains the apparent anomaly of the asymmetric terms being significant while the R^2adj for these models is less than for the corresponding symmetric models.[5] As expected, the coefficient on the threshold term is positive, implying that market downturns in a certain trading day are followed by increased volatility the next day. The leverage effect is also relatively large, suggesting that this effect plays a substantial role. Overall, table 1 shows that, of all the ARCH process models, the GARCH(2,2) provides the best approximation of the underlying volatility generating process, although this is marginal: all of the GARCH models have approximately the same forecasting ability.

The historical volatility models are not shown in table 1 as the relationships between them and the volatility are constructed, not estimated. With regards to their explanatory power, the correlation between the squared excess returns and the naïve forecast was 0,215, and with the squared excess returns and the moving average was 0,396. The moving average therefore provides a better fit than the naïve. For the SAVI, which is also not shown in the table as there are also no parameters to be estimated, the correlation with the squared excess returns is 0,368.

While the information contained in Table 1 provides evidence with respect to which of the volatility specifications best fit the domestic equity volatility in

[5]If the adjustments for increased explanatory terms was not made, the explanatory power increases only slightly.

sample, the next section focuses on which of the above specifications provides the best out of sample fit.

7. OUT OF SAMPLE FORECASTS

7.1 Forecast generation and evaluation

As the paper seeks to evaluate how well the various specifications forecast equity volatility they are generated using rolling windows. Forecasting the future volatility using the extrapolation models that have been estimated using information across the full sample will mean that the forecast for period T will include information regarding the actual at T. Forecasts generated in this fashion cannot be considered ‘forecasts’ in the true sense of the word. All of the extrapolation models’ forecasts of future volatility will therefore be out of sample forecasts, using the process of rolling estimation windows.

These rolling estimation windows were constructed as follows. Following Blair, Poon and Taylor (2001) the forecast of volatility made at period T the model was estimated across the previous 100 days. Once that forecast has been made, the model is re-estimated across a sample that now includes period T and the previous 99 days and the forecast made at period T+1 is generated. The process is repeated, with the model being re-estimated using a rolling window sample period of the previous 100 trading days. In this fashion, ‘true’ forecasts were generated for all the ARCH processes. As the historical volatility models do not include future information sets these rolling windows were not needed for the construction of their forecasts. This is arbitrarily true for the implied volatility forecast.

Two out-of-sample forecasts were generated, a one day ahead forecast and a one week (5 days) forecast. Following convention (Poon 2005), the volatility forecast for the one week period is taken as the sum of the individual forecasts, and evaluated against the actual daily volatility summed over those five days. In theory the forecast models should produce a better one week than a one day forecast due to the cancellation of forecast errors over the week and mean reversion of the actual volatility.

For the ARCH processes, the one week ahead forecasts were simple extrapolations of the model, with the forecasted values recursively substituted in at each step after T. These individual forecasts were then summed to gain the respective one week ARCH process forecast. For the naïve forecast, the 1 day ahead was multiplied by 5 to get the week forecast. For the 5 day MA forecast, rolling 1 day ahead forecasts were generated for each of the 5 days, using the calculated MA forecasts for the periods after T. These were then summed to gain the one week forecast. All of the 1 week forecasts then only use information available at time T, ensuring that they are all ‘true’ forecasts.

Moving on to the evaluation of the forecast ability, whether relative or absolute, it is obviously necessary to have some criteria that compares the costs and benefits of the forecasts. Evaluations that are based on some underlying utility function, such as that by West, Cho and Edison (1993), are not universally ideal as the shape and properties of the relevant utility function is not always known. In practise, most evaluations of competing forecasts use some measure of forecast fit that excludes any need for such utility assumptions, and simply attempts to assess the forecasts using measures based on the number, degree and direction of the forecast errors. This paper employs the widely used and well known non-utility base measures of the Mean Square Error (MSE), the Mean Absolute Error (MAE), Root Mean Square Error (RMSE) and the Mean Error (ME).

$$MSE = \frac{\sum (A_t - F_t)^2}{n} \quad \dots (11)$$

$$MAE = \frac{\sum |A_t - F_t)|}{n} \quad \dots (12)$$

$$RMSE = \sqrt{\frac{\sum (A_t - F_t)^2}{n}} \quad \dots (13)$$

$$ME = \frac{\sum (A_t - F_t)}{n} \quad \dots (14)$$

where

A_t is the actual volatility for period t, F_t is the forecasted future volatility and n is the number of periods over which the forecasts are evaluated. As the values of daily variance being compared are small in an absolute sense, for purely ease of read considerations this figures are scaled up by multiplying them by 10^6. Of these measures, note that the ME is clearly the inferior measure as its cancelling property will lead to inaccurate yet symmetric forecasts appearing as accurate predictors of the actual. The weight of interpretation should therefore always fall mainly on the other measures of fit. In addition, the major difference between the MAE and the MSE is that the MSE penalises large deviations from the actual more heavily than does the MAE.

In addition, for a forecast to be considered ‘good’ in an absolute sense it should at the very least not systematically under- or over- predict the level of volatility. Formally, for a one period ahead forecast this unbiased requirement of relationship between the forecasted variable and the actual is written as:

$$E[At - F_t \mid \Omega_{t-1}] = 0 \qquad \dots (15)$$

where A_t is the actual value of the variable known at time t, F_t is the forecast of the volatility for period t made at period $t\text{-}1$, and Ω_{t-1} the information set available to the market at time $t\text{-}1$. Operationally, the literature frequently tests the above formulation of bias in the forecasts by regressing the forecasted variable on the actual variable. This method, known as the Mincer-Zarnowitz regression, involves the estimation of the following model and the testing the joint unbiased hypothesis that α_F=0 and β_F=1:

$$A_t = \alpha_t + \beta_f F_t + \upsilon_t \qquad \dots (16)$$

The results of the estimation of equation (16) for all the models for the sample excluding the first 100 days is given in the tables below. Note that, due to problems of possible serial correlation in the estimation of equation (16) Newey-West Heteroscedasticity and Autocorrelation Consistent (HAC) standard errors were always calculated. The table includes the following information: the estimated coefficients for α_F and β_F, the Wald test of joint unbiasedness, the error evaluation figures (MSE/MAE/ME/RMSE) and the adjusted Coefficient of Determination (R^2_{Adj}).

Table 2: Out of sample forecast ability (10/05/2004 to 28/09/2006)

Panel A: 1 Day Ahead								
Model	α_F	β_F	MSE	MAE	RMSE	ME	Prob(Unbiased)	R^2_{adj}
ARCH	0,000*	0,735	0,0125	153,11	0,1118	5,1191	0,013	0,037
GARCH(1,1)	0,000	0,712	0,0115	157,81	0,1072	9,8150	0,449	0,130
GARCH(2,1)	0,000	0,683	0,0118	161,62	0,1086	13,6211	0,268	0,117
GARCH(1,2)	0,000	0,679	0,0118	156,14	0,1086	8,1448	0,351	0,115
GARCH(2,2)	0,000**	0,670**	0,0118	157,17	0,1086	9,1762	0,201	0,114
TARCH(1,1)	0,000	0,861	0,0106	146,19	0,1030	-1,8044	0,446	0,182
TARCH(2,2)	0,000	1,072	0,0103	148,45	0,1015	0,4525	0,968	0,197
SAVI	0,000	0,624*	0,0116	164,92	0,1077	16,9272	0,000	0,135
Naïve	0,000*	0,215*	0,0175	197,35	0,1323	49,3594	0,000	0,044
MA(5)	0,000*	0,625*	0,0114	153,41	0,1068	5,4174	0,025	0,155
Panel B: 1 Week Ahead								
Model	α_F	β_F	MSE	MAE	RMSE	ME	Prob(Unbiased)	R^2_{adj}
ARCH	0,000*	0,789*	1,2273	531,15	1,1078	-206,89	0,040	0,104
GARCH(1,1)	0,000*	0,581*	1,2101	537,86	1,1000	-200,18	0,000	0,217
GARCH(2,1)	0,000*	0,602*	1,1797	519,01	1,0861	-219,03	0,000	0,223
GARCH(1,2)	0,000*	0,577*	1,2174	556,66	1,1034	-181,39	0,000	0,217
GARCH(2,2)	0,000*	0,584*	1,2039	539,15	1,0972	-198,90	0,000	0,218
TARCH(1,1)	0,000*	0,713*	1,0651	488,93	1,0320	-249,12	0,000	0,251
TARCH(2,2)	0,000*	0,849	1,0158	470,44	1,0079	-267,60	0,094	0,253
SAVI	0,000*	0,557*	1,2554	579,83	1,1204	-158,21	0,000	0,234
Naïve	0,000*	0,262*	2,8808	786,35	1,6973	48,30	0,000	0,165
MA(5)	0,000*	0,515	1,4664	518,80	1,2110	-219,24	0,000	0,264

$\sigma^2_A = \alpha + \beta_F \sigma^2_F$
where
σ^2_A is the actual volatility measure, defined as Squared Excess Returns
σ^2_F is the forecasted value of the volatility

Notes:
MSE/MAE/RMSE/ME figures stated are actual multiplied by 10^6 (1000000)
Prob(Unbiased) is the joint probability that α_F=0 and β_F=1
Significance tests are for α = 0 and β_F = 1, respectively
The SAVI MSE/MAE/RMSE/ME figures are generated with α_F=0 and β_F=1 imposed.
*Significant at the 5% level
**Significant at the 10% level

7.2 One day ahead forecast

Looking at the absolute forecasting ability first, it is clear from the adjusted Coefficient of Determination that the ARCH forecast is a poor guide in respect to future volatility, explaining less than 4% of the volatility. The GARCH(1,1) specification provides the best forecast of all the symmetric specifications, although this is only marginal: all of the symmetric models forecast around 12% of the volatility. The asymmetric TARCH models provide far better forecasts of volatility than the symmetric in an absolute sense, with the TARCH(2,2) forecasting almost 20% of the equity volatility. The historical volatility models have a mixed out-of-sample absolute forecasting record. The naïve forecasts are poor, while the 5 day moving average forecasts are more accurate than all of the symmetric ARCH process models. This is a surprising result, and implies that the information concerning the underlying (symmetric) volatility structure does not add sufficient forecasting power to the day ahead forecast. The market (SAVI) one day ahead forecast provides a better guide to the future volatility (13,5%) than all of the symmetric ARCH process models (12%), though does not outperform the asymmetric (19%) or the 5 day MA forecast (15,5%). The MSE, MAE and RMSE measures provide further information supporting all of the conclusions drawn above. As higher values indicate poorer forecast records, the results drawn mirror that of the adjusted Coefficient of Determination. It appears that the most accurate one day ahead forecast is by the TARCH models, followed by the 5 day MA and then the market forecast.

Turning to the measures of bias, according to the joint Wald test of bias the only biased ARCH process forecast is that of the simple ARCH specification. Both of the historical volatility forecasts are found to be biased. Interestingly, the SAVI market forecast is also found to be biased. With regards to the actual coefficients, the beta coefficients show the bias in better detail. As estimated values greater than unity implies under-forecasting (and less than unity imply over-forecasting) it can be clearly seen than most of the forecasts over-predict domestic equity volatility. In the case of the naïve forecast and to some degree with the GARCH (0,577 – 0,602) and SAVI (0,557) forecasts, the estimate coefficients imply that this over-prediction is quite large. In contrast, the TARCH(2,2) specification, the most accurate forecast specification, shows evidence (1,072) of slight under-prediction of volatility. These patterns should not be overstressed though; the formal inference tests show that only the naïve, 5 day MA, the SAVI and the simple ARCH model have this over prediction bias as being a significant property.

7.3 One week ahead forecast

As was expected, the one week forecasts were more accurate than the one day ahead point forecasts. According to both the R^2_{Adj} and the MSE, all of the forecasts are substantially more accurate; note in particular the increased amount of volatility explained by the naïve and ARCH specifications. The general conclusions drawn from the 1 day ahead forecast regarding the relative forecasting record are still mostly the same. Simple ARCH is the poorest ARCH process forecast, the symmetric GARCH models have approximately the same forecasting power and the asymmetric GARCH models are more accurate than their symmetric counterparts.

An interesting change is that the 5 day moving average forecast, a relatively accurate 1 day forecast, is now more accurate than the all of the other forecasts according the Coefficient of Determination. Explaining over a fifth of the volatility over the full week, it clearly dominates all the symmetric GARCH process models, and explains slightly more than the asymmetric GARCH process models and the SAVI. Given it's significantly simpler estimation and manipulation this is a surprising result, though not a unique finding (Poon, 2005). This must however be qualified by looking at the error evaluation measures, MSE, MAE and the RMSE. According to these figures the asymmetric TARCH models are again the most accurate, and the SAVI is better (in terms of minimising errors) than the 5 day MA. It appears then that, as with the one day ahead forecast, the most accurate one week ahead forecast is by the TARCH models, followed by the market forecast and the 5 day MA forecast.

Looking at the bias tests, of all the forecasts only the TARCH(2,2) model provides evidence of not being biased. As with the one day ahead forecast, there is significant evidence of over-prediction of domestic volatility by the various specifications. Among the symmetric GARCH process models, the naïve forecast and the MA forecast, this is quite a significant effect. The beta coefficient of the SAVI forecast shows evidence of the market forecast over predicting volatility by a half. This should be borne in mind when using it. While the adjusted Coefficient of Determination can be used as a guide to infer which model is more accurate, this general result of over prediction must be noted by practitioners.

8. CONCLUSION

This paper investigated the forecasting of domestic equity volatility, a measure of risk facing both financial practitioners and theorists. Defining the volatility as being the square of the return in excess of the expected amount, the paper found that the ARCH process models provided a good in-sample approximation of the underlying domestic equity volatility process, with the exception of the simple ARCH specification. The historical models, naïve and 5 day MA, were found to be correlated in sample with the domestic equity volatility, as was the market forecast.

For the out of sample forecast ability, two forecasts were generated, a one day ahead and a one week ahead forecast. This paper found that simple ARCH and naïve forecasts are relatively quite poor predictors of volatility. The more complex GARCH specifications are better than the ARCH specifications and that the even more complex asymmetric models provide the best forecasts of all the specifications investigated. The TARCH(2,2) model is judged to be the best specification to model and forecast domestic volatility, being relatively the most accurate and unbiased over both one day and one week ahead forecast horizons. From an investors perspective, the use of asymmetric models provide better forecasts than the market does for a one day ahead forecast, while for a one week ahead forecast the SAVI forecast is little different from these asymmetric specifications. The 5 day MA forecast, despite being simple, provides a good forecast of volatility, rivalling the GARCH forecasts. The simple ARCH, the SAVI, the 5 day MA and the naïve forecast all over predict one day ahead domestic volatility, while all of the specifications except the TARCH(2,2) model over predict one week ahead volatility.

REFERENCES

Andersen T and Bollerslev T. 1998. Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review,* 39(4):885-903.

Blair B, Poon S and Taylor S. 2001. Forecasting S&P volatility: the incremental information content of implied volatilities and high-frequency returns. *Journal of Econometrics*, 105:5-26.

Bollerslev T and Wooldridge J. 1992. Quasi-maximum likelihood estimation and inference in dynamic models with time varying covariances. *Econometric Reviews*, 11:143-172.

Bollerslev T. 1986. Generalised autoregressive conditional heteroskedasticity. *Journal of Econometrics,* 31:307-327.

Bollerslev T, Chou R and Kroner K. 1992. ARCH modelling in finance. *Journal of Econometrics*, 52(1):5-59.

Cadiz. 2006. The new implied volatility index with applications. November.

Canina L and Figlewski S. 1993. The information content of implied volatility. *Review of Financial Studies*, 6:3 659 – 681.

Cubbin E, Eidne M, Firer C and Gilbert E. 2006. Mean reversion on the JSE Securities Exchange. *Investment Analysts Journal,* 63:39-47.

Engle R. 1982. Autoregressive conditional heteroskedasticity with estimates of the variance of UK inflation. *Econometrica*, 50:987–1008.

Markowitz H. 1952. Portfolio selection. *Journal of Finance*, 7(1):77-91.

Mincer J and Zarnowitz V. 1969. The evaluation of economic forecasts. In J. Mincer, ed,. *Economic Forecasts and Expectations*, New York: National Bureau of Economic Research, 1969.

Nelson D. 1989. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica,* 59:347-370.

Poon S. 2005. *A practical guide to forecasting Financial Market Volatility*. John Wiley and Sons, The Atrium, Chichester, West Sussex.

Tobin J. 1958. Liquidity preference as behaviour towards risk. *Review of Economic Studies*. 25(2):65-86.

West K, Cho D and Edison H. 1993. A utility based comparison of some methods of exchange rate volatility. *Journal of International Economics*, 35:23-45.

Whaley R, 2000. The investor fear gauge. *The Journal of Portfolio Management*, Spring.