The Information Content of Implied Volatility: Evidence from South Africa

Nico Katzke[a,b,*], Hanjo Odendaal[c], Albertus van Niekerk[d], Willem van Lill[d]

[a] *Department of Economics, Stellenbosch University, South Africa.*
[b] *Fairtree Capital, Willowbridge Place, Durbanville*
[c] *Eighty20 Analytics,145 Sir Lowry Road,Woodstock,Cape Town*
[d] *Econex,Stellenbosch,Cape Town*

Abstract

In this paper we investigate the predictive power of implied volatility as extracted from aggregated options prices traded on JSE ALSI top 40. Specifically, we investigate the information content of forward market volatility, as implied by the price of options, and test whether it adds meaningful forward information to predictions of realized volatility (RV) compared to several different backward looking second order persistence model estimates. We also test whether other proxies of domestic and global forward volatility, such as the SAVI and the CBOE VIX, provide meaningful information about RV.

Keywords:
Multivariate GARCH, Kalman Filter, Rolling window OLS, time-varying betas, South Africa sector analysis
JEL: L250, L100

---

1. Introduction

Investors place great value on the ability to accurately predict stock market volatility. It enables them to foresee risks and allows for the implementation of appropriate hedging strategies. Accordingly, a vast literature

*Corresponding author
Email addresses: nicokatzke@sun.ac.za (Nico Katzke), hanjo.odendaal@eighty20.co.za (Hanjo Odendaal), albertus@econex.co.za (Albertus van Niekerk), willemvl@econex.co.za (Willem van Lill)

attempts to accurately forecast volatility. Different models that have subsequently been developed include GARCH-type models (Bollerslev et al., 1992), stochastic volatility models (Taylor, 1986) and realised volatility models (Andersen et al., 1999), which rely on historical stock price data to forecast future volatility. However, another part of the literature relies on the informational content of stock options in predicting volatility. This entails using implied volatility (derived from asset pricing models) as a forecast of future volatility. Much of the recent research on the information content of implied volatilities employ a statistical technique that permits asset returns to vary over time according to a generalized autoregressive conditional heteroscedasticity (GARCH) model (see Bollerslev et al. (1992)). The GARCH class of models provide a more general framework for evaluating the incremental information content of implied volatility than, for example, a model relying on cross-sectional regressions. One major advantage of the GARCH models is that it explicitly accounts for the persistence of the volatility process and other time series features, as well as taking into account the relation between expected returns and conditional market volatility. We examine the predictive content of different implied volatility measures by adding them to the GARCH models as exogenous variables. In addition to the GARCH (1,1) model, we use a GJR -GARCH (1,1) model, which takes asymmetries into account. Through constructing a nested moel, we can statistically verify whether implied volatility is an important determinant of realized volatility. This also gives an indication of the efficiency of the options market.

This essay examines the information content of different measures of ex ante future market volatility implicit in the price of call options. Specifically, we use two domestic measures, implied volatility (IV) and the South African Volatility Index (SAVI), as well as a proxy for global market volatility, the Chicago Board Options Exchange's (CBOE) Volatility Index (VIX). Due to the fact that the GARCH models are fitted over the entire sample period, they characterize the within-sample properties of implied volatilities. However, these within-sample tests may bias the results in favour of the GARCH models, and as such we also test the out-of-sample forecasting power of implied volatility and forecasts from GARCH models. We find that implied volatility does in certain instances contain incremental information relative to model forecasts. However, when one is concerned with the accuracy of forecasting ex ante volatility, using both a measure based on market sentiment along with a measure based on the forecast from a statistical model yields the best results.

The remainder of this paper is organised as follows. Section 2 is a literature review on the informational content of implied volatility in United States (US) markets. Section 3 distinguishes between implied volatility, the SAVI and the VIX. Section 4 discusses the data used in the analysis and provides summary statistics. Section 5 elaborates on the methodology we use throughout the essay. Section 6 provides the in-sample forecasting results and also reports on the impact of the financial crisis. Section 7 discusses the out-of-sample forecasting performance of the models

2. Literature Review

The advent of the Black and Scholes (1973) options-pricing formula allows for the calculation of implied volatility, which "represents the market's estimate of the future volatility of the underlying asset for the lifetime of the option" (Frijns et al., 2010). Early empirical work indicates that forward-looking implied volatility is a better predictor of future volatility than historical standard deviations (Latane and Rendleman, 1976; Chiras and Manaster, 1978; Beckers, 1981). This forward-looking feature makes implied volatility a good measure to explain stock market returns (Merton, 1980). Studies by Christensen and Prabhala (1998), Fleming (1998), Blair et al. (2010) and Giot (2002) find that accurate volatility forecasts for stock market returns are very often based on implied volatility. The contemporaneous relationship between stock market returns and the VIX (Volatility Index) is negative and asymmetric (Fleming et al., 1995; Whaley, 1993). Consequently, increased implied volatility results in larger negative returns than the positive returns from decreased implied volatility. This feature allows it to function as a "fear gauge" for investors and reflects investor sentiment (Baker and Wurgler, 2006; Fleming, 1998). Implied volatility can also be used as a method for evaluating asset pricing models (Fleming, 1998). However, there is conflicting evidence concerning the ability of implied volatility to perform these roles efficiently.

Christensen and Prabhala (1998) use monthly data over an eleven and a half year period to test the hypothesis that the options on the S&P 100 index are efficiently priced. An options market will be efficient if implied volatility is an accurate forecast of future volatility. This suggests that implied volatility should subsume the information content in the other market variables in explaining future realised volatility (Christensen and Prabhala, 1998). They conclude that implied volatility adds incremental information to the forecast of conditional volatility and is a good predictor of ex-post realised volatility. The options on the S&P 100 index are therefore efficiently priced (Christensen and Prabhala, 1998). A large portion of the literature agrees with this point of view (see Poterba and Summers (1984); Fleming et al. (1995); Granger and Poon (2001); Mayhew and Stivers (2003); Becker et al. (2009).

Day and Lewis (1992) also examine the information content of implied volatility from call options on the S&P 100 index by adding implied volatility to GARCH and Exponential GARCH (EGARCH) models. They then test whether the coefficient on implied volatility comes out significant. Their in-sample results suggest that implied volatility may contain incremental information relative to the conditional volatility forecast generated by the GARCH and EGARCH models. However, their out-of-sample results prove to be inconclusive regarding the relative information content of GARCH forecasts and implied volatilities (Day and Lewis, 1992).

Canina and Figlewski (1993) also examine the options on the S&P 100 index and conclude that implied volatility is a poor forecast of realised volatility and therefore does not add incremental information to the forecast of volatility (Canina and Figlewski, 1993). Studies by Becker et al. (2007) and Becker et al. (2009) show that historical data subsumes information that is not incorporated into the stock price options, which suggests that implied volatility is a poor forecaster of volatility.

As a result, there is no clear consensus on the information content of implied volatility as opposed to historical volatility. From the studies mentioned above, there are conflicting views concerning the efficiency of the S&P 100 options market. This essay will conduct a similar analysis for South Africa by determining whether implied volatility of the JSE Top 40 is able to add information content to the volatility forecast of the underlying asset. This is also a measure of the efficiency of the market. Furthermore, this essay compares different implied volatility measures and also evaluates their performance with specific reference to the financial crisis.

3. Different measures of Implied Volatility

As mentioned earlier, implied volatility is derived from options-pricing formulas like the Black-Scholes. All the variables of the formula are known. The only unknown variable is volatility. Manipulation of this formula allows for the calculation of implied volatility using the known variables as inputs.

This represents the market's forecast of volatility over the lifetime of the option and is therefore regarded as the market's "fear gauge" (Baker and Wurgler, 2006; Fleming, 1998).

Implied volatility has a non-linear trend of implied volatility at different strike prices (Poon & Granger, 2003). This non-linearity is known as the volatility smile, smirk, or sneer. As a result, empirical evidence suggests that risky financial assets have leptokurtic tails. When the strike price is very high, the option is deep out-of-the-money with a low probability of being exercised. A leptokurtic right tail gives a higher probability for the terminal asset price to exceed the strike price and the call option to finish in-the-money. This results is a higher strike price and higher Black-Scholes implied volatility (Granger and Poon, 2001).

Many weighting schemes have been proposed to overcome the bias imposed by the volatility smile (Bates, 1996). However, since the plot of implied volatility against the strike price can take a variety of shapes, no single weighting scheme can consistently remove pricing errors (Granger and Poon, 2001). To overcome the bias of the volatility smile, this study uses at-the-money (ATM) implied volatility of options that have 90 days to expiration. Several other measures of implied volatility exist. For the purpose of this essay, we will focus on the SAVI and the VIX.

3.1. South African Volatility Index (SAVI)

Volatility reflects the risk in financial markets. It can provide an estimate or forecast of how far prices are expected to move in a given time (Kotz? et al., 2009). Lower volatility indicates complacency in the markets, whereas higher volatility reflects a fearful market. Investor behaviour is driven by uncertainty. Market sentiment is therefore crucial in making investment decisions. The South African Volatility Index (SAVI) is the South African "fear gauge" and is constructed as the weighted average prices of calls and puts over a wide range of strike prices that expire in three months (Kotz? et al., 2009). It is therefore an index of the market's expectation of the 3-month market volatility.

The SAVI is calculated as:

$$SAVI = \sqrt{\sum_{i=1}^{n=F} w_{ip} P_i(K_i) + \sum_{i=n}^{\infty} w_{iC} C_i(K_i)} \quad (1)$$

Where F is the current (on value-date) forward of the FTSE/JSE Top 40 index level . F is the price boundary between the liquid put options, $P_i(K_i)$, and call options, $C_i(K_i)$, with strikes K_i. Call and put option prices are calculated using the traded market volatility skew that has three months (90 days) to expiration. The 3-month volatility skew is determined using a time weighted interpolation function. The weightings are piecewise linear recurring option weightings and are denoted by w_{iP} and w_{iC} (Derman et al., 1999). A strike spacing of 10 index points leads to insignificant approximation errors within a 70% to 130% moneyness range (Kotz? et al., 2009). Therefore the South African Futures Exchange (Safex) uses this 70% to 130% option moneyness range to calculate the SAVI, while correcting for the volatility skew (Kotz? et al., 2009).

3.2. Chicago Board Options Exchange Volatility Index (VIX)

The VIX was created by the Chicago Board Options Exchange (CBOE) in 1993. The initial VIX was constructed using near-the-money options on the S&P 100 index. These options have a constant time to maturity of 30 days. The VIX calculation was adapted in September 2003 to use bid and ask quotes of the S&P 500, the benchmark index in the US (Huang, 2011). The VIX is the market's expected volatility of the S&P 500 index over the next 30 days. It is computed directly from option prices rather than being derived from the Black-Scholes option pricing formula. Using at-the-money and out-of-the-money call and put options, it is calculated using the formula:

$$\sigma^2 = \frac{2}{T}\Sigma\frac{\Delta K_i}{K_i^2}e^{rT}Q(K_i) - \frac{1}{T}\left[\frac{F}{K_0} - 1\right]^2 \quad (2)$$

where σ is the VIX divided by 100, T is the time to maturity, r is the risk-free interest rate, F is the forward index level calculated by index option prices. K_0 is the first strike below $F; K_i$ is the i^th out-of-the-money option. $Q(K_i)$ is the midpoint of the bid-ask spread for each option with strike K_i (Huang, 2011). ΔK_i is the interval between strike prices and is computed as the difference between the strikes on either side of $K_i : \frac{(K_{i+1} - K_{i-1})}{2}$

The near-term options of the first two contract months are used in calculating the VIX. Consequently, for these options the time to maturity should be at least one week in order to minimize potential pricing anomalies that can occur near the expiration of the option (Huang, 2011).

4. Data

The data used in the study are daily closing prices from the JSE Top 40 index as well as three different implied volatility measures, as outlined in the previous section. The continuously compounded daily returns are calculated by taking the log difference of the index:

$$r_t = ln(\frac{p_t}{p_{t-1}}) * 100 \tag{3}$$

with p_t the closing price of the Top 40 Index at time t. Table (1) reports the summary statistics of both full-sample and annual sub-sample periods

	T	μ	σ	$\tilde{x}$	Max	Min	Skew	Kurtosis	JB	Prob
All	1775.00	0.04	1.48	0.10	7.71	-7.96	-0.08	6.24	781.00	0.00
2007	238.00	0.06	1.32	0.20	3.95	-4.51	-0.45	3.93	17.21	0.00
2008	262.00	-0.12	2.46	-0.16	7.71	-7.96	0.07	3.99	11.52	0.00
2009	261.00	0.11	1.68	0.17	6.21	-3.89	0.14	3.35	2.45	0.29
2010	261.00	0.06	1.15	0.08	4.48	-3.84	-0.06	4.22	17.24	0.00
2011	260.00	-0.01	1.29	0.01	4.05	-3.44	0.02	3.52	3.28	0.19
2012	261.00	0.09	0.80	0.14	2.75	-2.65	-0.20	3.50	4.82	0.09
2013	232.00	0.08	1.00	0.06	2.53	-3.43	-0.45	4.10	20.49	0.00

Table 1: Descriptive statistics: Top40

The average return for the Top40 returns for the period ranges between -7.96% to 7.71% over the period 2007 - 2013, showing a mean return of 0.037%. The 2008 financial crisis is also prevelent in the distribution of returns with a yearly mean return of -0.118% accompanied by the largest standard deviation in the sample of 2.455.

The data also conforms to the stylized facts of daily compounded return day. The leptocurtic distribution of the returns are confirmed with the Jarques-bera test for normality being rejected for all years excluding 2008 and 2011. This deviance from non-normality is usually addressed in the form of applying a fat-tailed distribution such as the Student-t, or any of the mixture models to a series in estimation.

Figure (1) represents both the daily compounded return series as well as the squared return series for the data. The squared return series is of interest as it serve as a proxy for realized volatility for the period (Frijns et al., 2010). It also appears that when evaluating the squared returns, a pattern of volatility clustering is observed at times. This is characterised

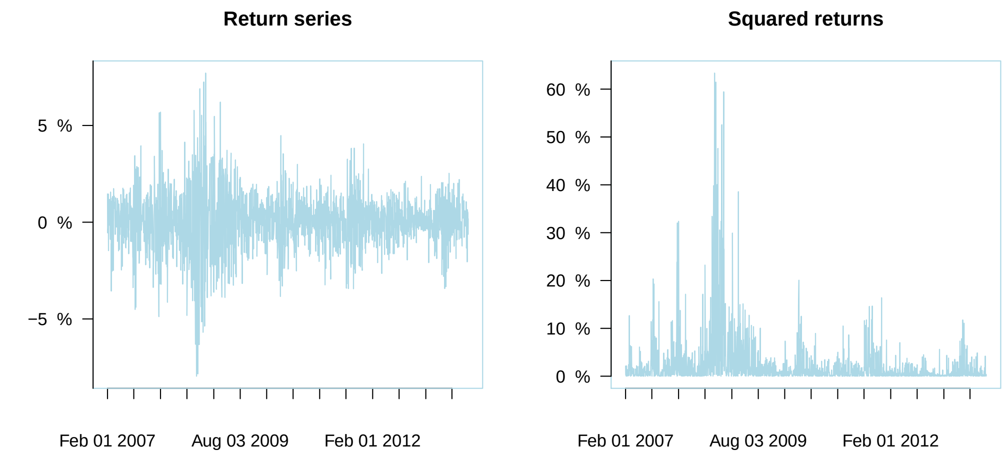

Figure 1: Top 40 daily returns

by periods of relatively small return movement, followed by volatile periods whereby the return series does shift relatively largely. It was Mandelbrot (1963) and Fama (1965) who first documented this phenomenon of clustering within fat-tailed return distributions.

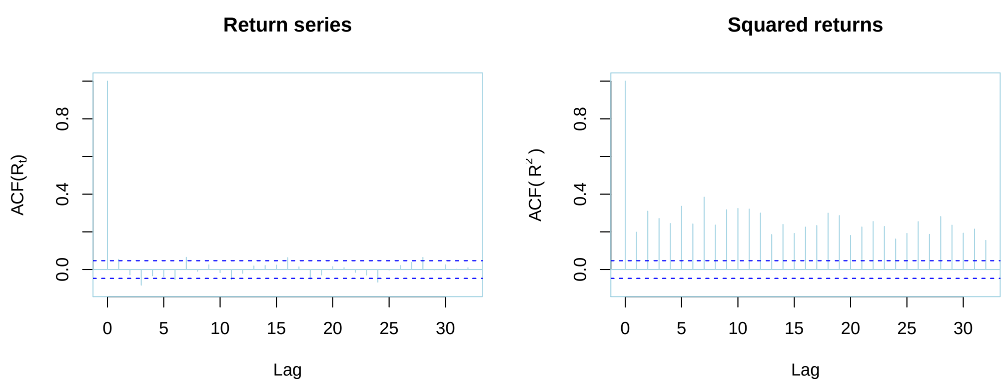

Figure 2: Top 40 daily returns

A visual representative of this phenomonon is presented in figure (2). Although the return series' autocorrelation graph shows little activity, the squared returns significant correlation with its past values for an extended lag period. The Box-Ljung statistic in Table (2) confirms the correlation for the squared returns series or so-called ARCH-effect for the sample in its entirety.

	Qsq: Return	Prob	Qsq: Squared Return	Prob
All	5.19	0.02	69.56	0.00
2007	0.51	0.98	0.60	0.00
2008	0.48	0.86	0.44	0.03
2009	0.71	0.35	3.12	0.86
2010	0.40	0.25	0.08	0.58
2011	3.74	0.62	0.15	0.45
2012	0.05	0.34	0.70	0.10
2013	0.00	0.56	13.91	0.75

Table 2: Box-Ljung Statistics: Top40

The domestic implied volatility series are the 100%-moneyness implied volatility measure and the SAVI. As a global measure of implied volatility, we use the VIX. Note that for the IV and SAVI, the implied volatility is derived from Top 40 options which have three months to expiration. While the SAVI and the VIX are computationally equivalent, the implied volatility used in the calculation of the VIX is derived from options that have one month to maturity. All implied volatility measures are reported in annualised terms, and as such we follow the consensus adopted in the literature in constructing series that are de-annualised . We do this by adjusting the annualised measure to take account of the amount of trading days in a year:

$$\theta_t^{daily} = \frac{\theta_t^{annual}}{\sqrt{252}} \tag{4}$$

where θ denotes the specific implied volatility measure ($\theta \in \{IV_t; SAVI_t; VIX_t\}$)

The return series, as well as the SAVI series, is obtained from McGregor BFA, whereas the IV data is obtained from Bloomberg. The VIX data is obtained from the CBOE. The in-sample data spans from 1 February 2007 to 20 November 2013. To obtain out-of-sample estimates, we use the return data from 1 February 2007 to 3 June 2013 and then apply a rolling forecast till the end of the sample. More details will follow in the next section. In total, the in-sample period contains 1775 observations, while the out-of-sample period contains 95 observations.

Tables (3)-(5) below provide summary statistics for the different measures of implied volatility for both full-sample and annual sub-sample periods.

	T	μ	σ	$\tilde{x}$	Max	Min	Skew	Kurtosis	JB	Prob
All	1775.00	24.62	7.35	23.09	57.97	13.51	1.51	6.13	1404.62	0.00
2007	238.00	22.41	3.40	21.67	31.07	16.46	0.32	2.15	11.11	0.00
2008	262.00	33.41	9.25	30.43	57.97	20.63	1.00	2.85	44.64	0.00
2009	261.00	29.54	7.11	27.72	47.69	20.33	0.73	2.27	28.90	0.00
2010	261.00	24.45	3.33	23.82	33.13	18.62	0.45	2.38	12.87	0.00
2011	260.00	24.90	3.37	23.73	34.07	20.43	0.61	2.29	21.62	0.00
2012	261.00	19.68	3.08	20.53	26.37	13.63	-0.14	2.06	10.21	0.01
2013	232.00	16.85	2.15	16.45	22.97	13.51	1.11	3.62	52.80	0.00

Table 3: SAVI

	T	μ	σ	$\tilde{x}$	Max	Min	Skew	Kurtosis	JB	Prob
All	1775.00	23.21	10.55	20.47	80.86	10.02	2.04	8.13	3183.01	0.00
2007	238.00	18.07	5.23	17.03	31.09	10.02	0.42	2.13	14.50	0.00
2008	262.00	32.59	16.23	25.18	80.86	16.30	1.29	3.31	74.83	0.00
2009	261.00	31.55	9.12	28.68	56.65	19.47	0.67	2.17	26.78	0.00
2010	261.00	22.55	5.28	21.72	45.79	15.45	1.26	4.82	106.88	0.00
2011	260.00	24.12	8.15	20.68	48.00	14.62	0.76	2.36	29.73	0.00
2012	261.00	17.84	2.54	17.58	26.66	13.45	0.83	3.48	32.82	0.00
2013	232.00	14.27	1.79	13.74	20.49	11.30	1.17	4.09	65.98	0.00

Table 4: VIX

The series display an average implied volatility over the period of 24% with definite upward deviations between 2008 and 2009. The impact of the financial crises is clearly visible in all through measures, having the implied volatility measure being push above 30% for the crises years 2008 and 2009. During this period the ranges of volatility also increased with the VIX showing the greatest dispersion of implied volatility with a standard deviation measure of 16,23. This is also reflected in its distributional properties having a relatively high kurtosis when compared to the other implied volatility measures. This is to do with the VIX measure reflecting the turmoil conditions in the US after the onset of the financial crises. Figure (3) illustrates the relationship between the each series visually.

	T	μ	σ	$\tilde{x}$	Max	Min	Skew	Kurtosis	JB	Prob
All	1775.00	24.62	7.75	22.86	57.47	12.75	1.28	5.15	831.51	0.00
2007	238.00	24.57	4.08	23.58	32.32	17.38	0.06	1.64	18.04	0.00
2008	262.00	34.02	9.11	30.55	57.47	23.26	0.90	2.59	37.05	0.00
2009	261.00	30.17	7.09	28.48	46.55	20.99	0.61	2.05	26.17	0.00
2010	261.00	24.01	3.45	23.24	32.54	19.20	0.60	2.31	20.56	0.00
2011	260.00	23.87	3.72	22.50	36.31	19.05	0.81	2.88	29.20	0.00
2012	261.00	18.48	3.00	19.15	25.22	12.82	-0.02	2.10	8.61	0.01
2013	232.00	16.21	2.27	15.49	22.18	12.75	1.07	3.23	45.26	0.00

Table 5: IV

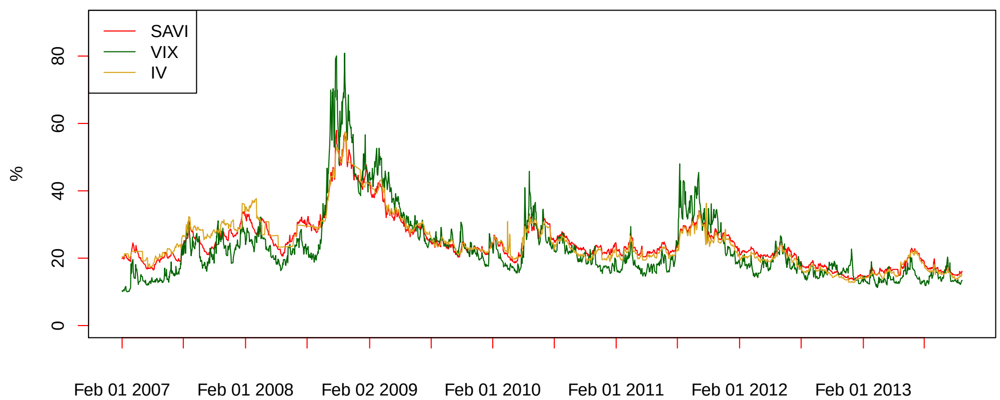

Figure 3: Implied volatility measures

A second striking feature from the figure is that prior to the crisis (which is shaded in the figure) there is a clear difference between the levels of domestic implied volatility and global implied volatility. However, following the crisis, the gap between the domestic and global measures of implied volatility has contracted and they track each other much closer. This may indicate that the South African economy has become much more susceptible to movements in the US economy during and after the crisis. This is especially evident in light of large spillovers from unconventional monetary policy measures implemented in the US. The crisis has therefore strengthened the interconnectedness between economies. There are noticeable peaks in the implied volatility measures over the sample period. The largest spike coincides with the Lehman Brothers bankruptcy and subsequent collapse of the interbank credit market. The first dotted line is at May 20, 2010, when concerns arose regarding the European debt crisis (Pepitone, 2010). The second dotted line is on August 8, 2011, when Standard & Poor's downgraded the US credit rating from AAA to AA+ (Pepitone, 2011). The smaller rise in the VIX between the second and third dotted line resulted from a dismal US jobs report and fears about the health of the global economy (Cox, 2012). The third dotted line is on December 28, 2012, and shows the market's growing concerns regarding the US fiscal cliff. On this day, the government had only three days left to reach a budget agreement to avert year-end tax increases and spending cuts (Kiernan, 2012). The fear gauge function of the implied volatility measures is showcased by the clear response to important news events. Figure (4) plots the JSE Top 40 index daily returns (rather than the level) along with the annualised percentage of the implied volatility measures.

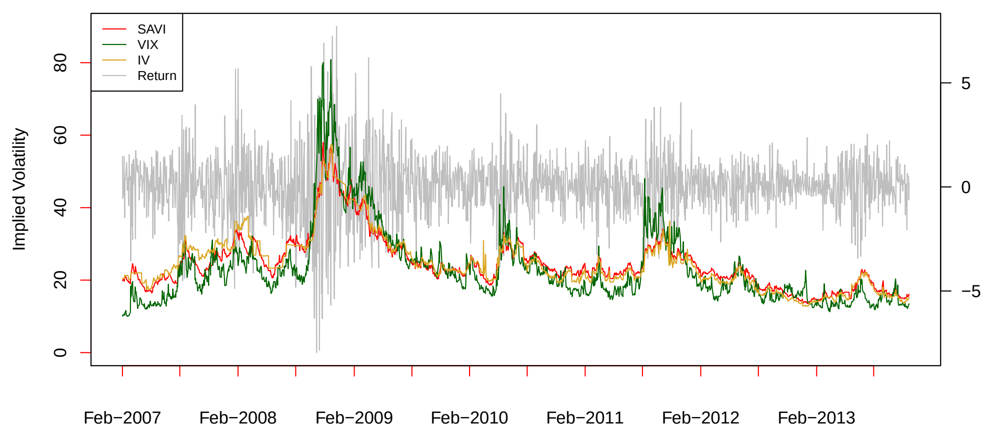

Figure 4: Implied volatility measures along return

In some periods, the positive returns associated with a decrease in implied volatility are smaller than the negative returns associated with an increase. As such, the figure provides evidence for the negative and asymmetric relationship that exists between implied volatility and returns. Finally, Figure 5 plots the absolute daily returns along with the annualised percentage of the implied volatility measures.

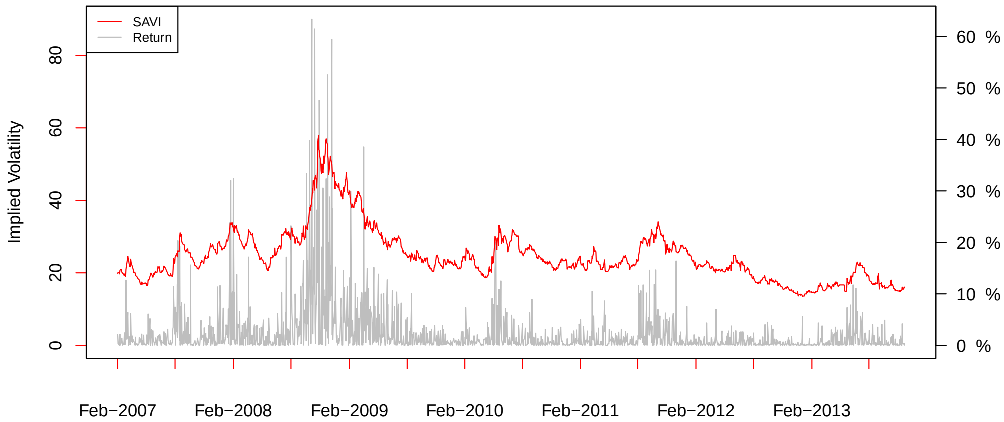

Figure 5: Top 40 daily absolute returns and SAVI series

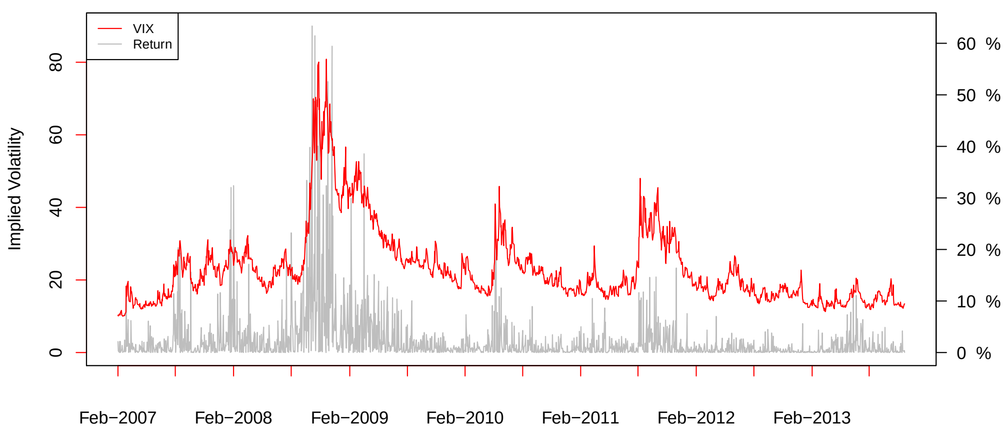

Figure 6: Top 40 daily absolute returns and VIX series

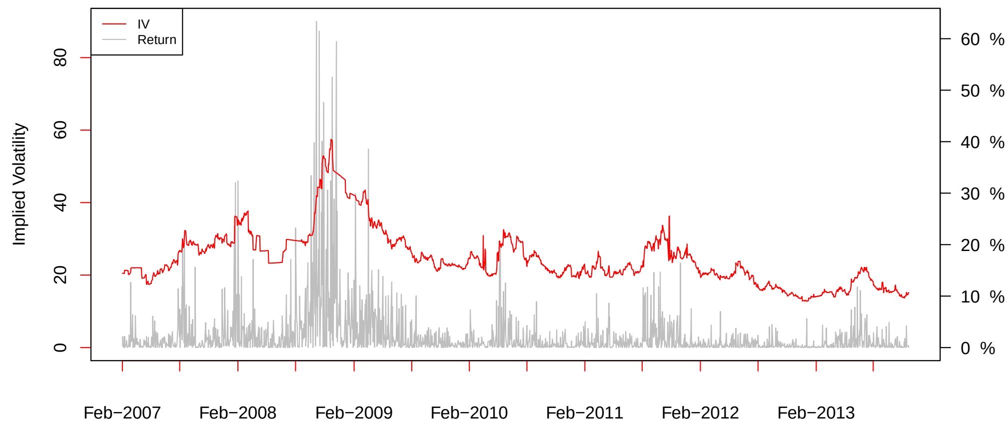

Figure 7: Top 40 daily absolute returns and IV series

It is clear from the figure that there is a close relationship between absolute daily returns and the implied volatility measures. As such, the implied volatility measures do well to capture market volatility. However, while the graphical analysis indicates that implied volatility is efficient at capturing market volatility, it is not clear which forecaster (IV, SAVI or VIX) performs better compared to the other, or compared to forecasts from econometric models. Also, we are not able to derive from the figures whether or not forecasters act in a complementary manner; namely whether using the IV in conjunction with, for example, the VIX results in better estimates than using either on their own. Taken together, the above implies the need to econometrically study the intertemporal relationship between different implied volatility measures and the continuously compounded daily returns from the JSE Top 40 index.

5. Methodology

In this section the methodological approach to determining whether information content from implied volatility measures statistically contribute to better volatility estimation is discussed. In the univariate case of GARCH models, the models themselve consists out of 2 equations. These are known as the mean equation which tries to model the evolution of the return series given a set of input variables plus an error term. In the second equation, the aim of the specification is to model the conditional variance of the error from its mean. It does by specifiying the variance as a function of past conditional variance and lagged errors. Although the mean equation is an important factor, in this paper only the variance equation will be evaluated.

5.1. The variance equation

In the mean equation the series y_t is modeled as a function of its mean and and error: $y_t = \mu_t + \varepsilon_t$. Here the mean, $\mu_t = E\{y_t|\mathcal{F}_{t-1}\}$ is modeled as the expected value given the information set $\mathcal{F}_{t-1} = \{y_{t-j}, j \geq 1\}$ in period $t-1$. The error, ε_t is independent and identically distributed, $\varepsilon_t \sim N(0, \sigma^2)$.

The standard GARCH(p, q) paremetrizes the error term into a deterministic, h_t and stochastic component, z_t as follows (Bollerslev, 1986):

$$\varepsilon_t = z_t \sqrt{h_t} \tag{5}$$

with $z_t \sim i.i.d(0, 1)$. The positive-valued conditional variance equaiton formulated is as:

$$h_t = \left(\omega_0 + \sum_{j=1}^{m} \zeta_j \upsilon_{jt}\right) + \sum_{j=1}^{q} \alpha_j \varepsilon_{t-j}^2 + \sum_{j=1}^{p} \beta_j h_{t-j} \tag{6}$$

with h_t denoting the conditional variance, ω the intercept,ζ_j the significance parameter of external variables, α_j the news effect, β_j the momentum carried forward from past variance and ε_t the residuals. As discussed in section (4), one of the key characteristics of financial data is its volatility clustering behaviours. The GARCH class of models are design to capture this through the peristence parameter $\hat{P}$. In the standard GARCH this is calculated as:

$$\hat{P} = \sum_{j=1}^{q} \alpha_j + \sum_{j=1}^{p} \beta_j \tag{7}$$

Another metric that relates to the peristence parameter is the "half-life" measure which captures the amount it takes in days for the volatility to revert back to $E(h_t)$:

$$h2l = \frac{-log_e 2}{log_e \hat{P}} \tag{8}$$

In a final discussion, the unconditional variance of the is related to its persistence parameter as:

$$\hat{h} = \frac{\hat{\omega}}{1 - \hat{P}} \tag{9}$$

5.2. The GJR-GARCH

In the model of Glosten et al. (1993), the variance formulation addresses the shortcoming of symetric interaction in the standard GARCH. It models the positive and negative news effects, α_j on the conditional variance asymmetrically through the use of an indicator function I.

$$h_t = \left(\omega_0 + \sum_{j=1}^{m} \zeta_j \upsilon_{jt}\right) + \sum_{j=1}^{q} \left(\alpha_j \varepsilon_{t-j}^2 + \gamma_j I_{t-j} \varepsilon_{t-j}^2\right) + \sum_{j=1}^{p} \beta_j h_{t-j} \tag{10}$$

The parameter, γ_j now serves as the "leverage" term, correcting for a bias that lies within negative residuals ε_t on variance. It does this by setting the indicator function $I = 1$ when $\varepsilon \leq 0$, and 0 otherwise. Along with the introduction of the indicator function, the persistence parameter also now adjusted to accommodate the asymmetry:

$$\hat{P} = \sum_{j=1}^{q} \alpha_j + \sum_{j=1}^{p} \beta_j + \sum_{j=1}^{q} \gamma_j \kappa \tag{11}$$

where κ represents the expected value of the standardized residuals z_t below zero. In essence it measures the probability of z_t being blow zero.

$$\kappa = E[I_{t-j} z_{t-j}^2] = \int_{\infty}^{0} f(z, 0, 1...) dz \tag{12}$$

where f is the standardized conditional density along with any additional skew and shape parameters. Given a symmetric distribution, the value of κ takes on 0.5.

5.3. Family of GARCH model

In extending the notation of modelling the asymmetric impact of the news-effect parameter on conditional variance, we can also estimate the so-called Taylor effect. Taylor (1986) observed that when looking at the autocorrelation of absolute and squared returns, that the autocorrelation for the absolute returns were much higher that of its squared counterpart.

One model that that acts as an omnibus of nested GARCH models is that of Hentschel (1995). This family GARCH model allows for the decompossed residuals in the conditional variance equation to be driven by different powers for z_t and h_t. It also allows for the integration of shift and rotation in the news impact curve. The shifting of the curve reflects the asymmetry for small shocks, while the rotation drives the larger effect of big shocks:

$$h_t^\delta = \left(\omega_0 + \sum_{j=1}^{m} \zeta_j \upsilon_{jt}\right) + \sum_{j=1}^{q} \alpha_j h_{t-j}^\lambda (|z_{i-j} - \eta_{2j}| - \eta_{1j}(z_{t-j} - \eta_{2j}))^\delta + \sum_{j=1}^{p} \beta_j h_{t-j}^\lambda \tag{13}$$

Eq (13) represents a Box-Cox transformation for the shape of the conditional standard deviation as determined by λ. Generally, for $\lambda > 1$, h_t is convex, while for $\lambda < 1$, it is concave The parameter δ transform the absolute value function which in turn is subject to rotation and shift through the η_{1j} and η_{2j} parameters.

Within this nested formulations various submodel specifications can be estimated. In this paper, we will however only focus on 2 of these. To relate the eq (13) to the standard GARCH in (6), the parameters are set to $\lambda = \delta = 2$ and $\eta_{1j} = \eta_{2j} = 0$. The first of these submodels that will be employed in the analysis will be the Absolute Value (AVGARCH) model of Taylor (1986) and Engle (1990). In the AVGARCH model the parameters are set to $\lambda = \delta = 1$ and $\eta_{1j} \leq 1$.

The second specification is the Nonlinear Asymmetric GARCH model of Engle and Ng (1993a). In this formultion the parameter η_{1j} is set to zero, with η_{2j} free to be estimated. This results in the news impact curve being able to shift, but not tilt, while retaining non-linearity by setting $\lambda = \delta = 2$.

The persistence parameter of the family GARCH model is given by:

$$\hat{P} = \sum_{j=1}^{q} \alpha_j \kappa_j + \sum_{j=1}^{p} \beta_j \tag{14}$$

now κ is the expected value of the standardized residuals, z_t given the Box-Cox transformation of the absolute value asymmetric term:

$$\begin{aligned}\kappa_j &= E(|z_{t-j} - \eta_{2j}| - \eta_{1j}(z_{t-j} - \eta_{2j}))^\delta \\ &= \int_{-\infty}^{\infty} (|z_{t-j} - \eta_{2j}| - \eta_{1j}(z_{t-j} - \eta_{2j}))^\delta f(z, 0, 1, ...)dz\end{aligned} \tag{15}$$

5.4. Introducing long-memory: Component GARCH

In order to capture the long memory component of volatility, Engle and Lee (1999) suggests decomposing the conditional variance into two components, namely its permanent and trasitory component. This enables the estimation of both the long- and short-run movements of volatility within the securitis market. The C-GARCH can be seen as an extension to the standard GARCH model with h_t being mean-reverting to a long term trend q_t, instead of a fixed position. In denoting q_t as the permanent component of the unconditional variance, the component model can be written as:

$$\begin{aligned} h_t &= q_t + \sum_{j=1}^{q} \alpha_j(\varepsilon_{t-j}^2 - q_{t-j}) + \sum_{j=1}^{p} \beta_j(h_{t-j} - q_{t-j}) \\ q_t &= \omega + \rho q_{t-1} + \phi(\varepsilon_{t-1}^2 - h_{t-1}) \end{aligned} \tag{16}$$

Effectively, what this formulation has accomplished is transforming the intecept into a time-varying first order autoregressive type dynamic. The transitory part of the volatiliy dynamic is thus formulated as being the difference between the conditional variance and its long-memory trend, $h_{t-j}^2 - q_{t-j}$. It is important to take note of the stationarity conditions, $\alpha + \beta < 1; \rho < 1$.

6. Conditional distributions

One of the key requirements of any autoregressive type process is that is should be self-decomposable while retaining the linear transformation property to center $(x_t - \mu_t)$ and scale (ε_t/σ_t). Thereafter can proceed using the zero-mean, unit variance dsitribution of the standardized innovation z_t. In this section the commonly known fact of fat-tail distribution of financial time series data is addresed in adjusting for this phenomenon in the conditional distribution of the innovation, z_t. In addressing this, formulation of gaussian innovation is replaced by more favoured fat-tailed distributions such as Student-t's distribution (std-t), the generalized error distribution (ged) and the Normal Inverse Gaussian (nig) as described in Ghalanos (2014).

6.1. Normal Distribution

The normal distribution is the most commonly used distribution in statistics and econometrics. The spherical distribution is completely described by its first 2 moments, mean (μ) and variance (σ^2). Assume a random variable, x, it can be said that it is normally distributed with density:

$$f(x) = \frac{e^{\frac{-0.5(x-\mu)^2}{\sigma^2}}}{\sigma\sqrt{2\pi}} \tag{17}$$

Using a mean filtering or whitening process, the standardized residuals, ε, yields the standard normal density:

$$f\left(\frac{x-\mu}{\sigma}\right)=\frac{1}{\sigma}f\left(z\right)=\frac{1}{\sigma}\left(\frac{e^{-0.5z^2}}{\sqrt{2\pi}}\right) \tag{18}$$

In obtaining the conditional likelihood of the GARCH process at each of the points (LL_t), the conditional standard deviation, (σ) from the GARCH motion dynamics acts as a scaling factor on the density:

$$LL_t\left(z_t;\sigma_t\right)=\frac{1}{\sigma_t}f\left(z_t\right) \tag{19}$$

Eq (19) thus illustrates the importance of a correct scaling factor in estimation.

6.2. The Student Distribution

In the paper of Bollerslev (1987), the Student-t distribution was first proposed as an alternative for fitting the standardized innovations. This was in response to the Normal distribution whose formulation lacked theoretical grounding. It can be described in its entirety through its degree of freedom, or shape parameter ν. In standardization, a 3 paramter representation is used:

$$f\left(x\right)=\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\beta\nu\pi}\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{\left(x-\alpha\right)^2}{\beta\nu}\right)^{-\left(\frac{\nu+1}{2}\right)} \tag{20}$$

where α, β, and ν are the location, scale and shape parameters respectively, while Γ is representative of the Gamma function. It relates to the normal distribution by having the same unimodal and symmetric shape, but with fatter tails. The location parameter, α is the mean and variance us defined as:

$$Var\left(x\right)=\frac{\beta\nu}{\left(\nu-2\right)} \tag{21}$$

A requirement of standardization is that the variance be declared as:

$$\begin{aligned} Var(x)&=\frac{\beta\nu}{\left(\nu-2\right)}=1\\ \therefore\beta&=\frac{\nu-2}{\nu} \end{aligned} \tag{22}$$

substituting the variance back into the 3 parameterized formulation we obtain the standardized Student's distribution:

$$f\left(\frac{x-\mu}{\sigma}\right)=\frac{1}{\sigma}f\left(z\right)=\frac{1}{\sigma}\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\left(\nu-2\right)\pi}\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{z^2}{\left(\nu-2\right)}\right)^{-\left(\frac{\nu+1}{2}\right)} \tag{23}$$

6.3. The Generalized Error and normal-inverse Gaussian Distribution

The distribution mentioned in subsection (6.2) was a special case of the also unimodal and symmetric generalized error distribution. The 3-parameter GED is often favoured due to its ability to capture the market dynamics through its semi-heavy tails. Its conditional density is given by:

$$f\left(x\right)=\frac{\kappa e^{-0.5\left|\frac{x-\alpha}{\beta}\right|^{\kappa}}}{2^{1+\kappa^{-1}}\beta\Gamma\left(\kappa^{-1}\right)} \tag{24}$$

where α, β and κ represents the location, scale and shape parameters. The standardized GED is then represented by:

$$f\left(\frac{x-\mu}{\sigma}\right)=\frac{1}{\sigma}f\left(z\right)=\frac{1}{\sigma}\frac{\kappa e^{-0.5\left|\sqrt{2^{-2/\kappa}\frac{\Gamma(\kappa^{-1})}{\Gamma(3\kappa^{-1})}}z\right|^{\kappa}}}{\sqrt{2^{-2/\kappa}\frac{\Gamma(\kappa^{-1})}{\Gamma(3\kappa^{-1})}}2^{1+\kappa^{-1}}\Gamma\left(\kappa^{-1}\right)} \tag{25}$$

For the NIG distribution, the formulation changes to:

$$X\sim \mathrm{GH}(-1/2,\alpha,\beta,\delta,\mu) \tag{26}$$

7. In-sample forecast evaluation

For in-sample evalution of the different volatility measures, we employ a likelihood ratio test. All models are tested using the simple model without the implied volatility measure, against the general model which has one more parameter due to the inclusion of ($\theta\in\{IV_t;SAVI_t;VIX_t\}$). As the names suggets, it is a ratio of two likelihood functions:

$$LRT=-2log_e\left(\frac{\mathscr{L}_s(\hat{\theta})}{\mathscr{L}_g(\hat{\theta})}\right) \tag{27}$$

where $\mathscr{L}_s$ deontes the simple model and $\mathscr{L}_g$ the general model. The test statistic is χ^2 distributed with the null hypothesis being that there is no difference in the accuracy of the two models. A non-rejection of the null hypothesis would thus imply that for in-sample testing the implied volatility measure had no information content to add. Table (6) represents the findings for the in-sample evaluation.

Table 6: In-sample results

	IV			SAVI			VIX			ALL		
	ζ_j	ζ_j sig	LRT	ζ_j	ζ_j sig	LRT	ζ_j	ζ_j sig	LRT	Average ζ_j		
AVGARCHged	0.0009	No	No	0.0001	No	No	0.0001	No	No	0.0004	0	0
AVGARCHnig	0.0006	No	No	0.0001	No	No	0.0003	No	No	0.0003	0	0
AVGARCHnorm	0.0007	No	No	0.0000	No	No	0.0000	No	No	0.0002	0	0
AVGARCHstd	0.0014	No	No	0.0003	No	No	0.0002	No	No	0.0006	0	0
csGARCHged	0.0000	No	No	0.0000	No	No	0.0000	No	No	0.0000	0	0
csGARCHnig	0.0000	No	No	0.0000	No	No	0.0000	No	No	0.0000	0	0
csGARCHnorm	0.1346	Yes	Yes	0.0000	No	No	0.0000	No	No	0.0449	1	1
csGARCHstd	0.0769	Yes	Yes	0.1931	Yes	Yes	0.0000	No	No	0.0900	2	2
NAGARCHged	0.0015	No	Yes	0.0006	No	No	0.0004	No	No	0.0008	0	1
NAGARCHnig	0.0013	No	No	0.0004	No	No	0.0007	No	No	0.0008	0	0
NAGARCHnorm	0.0016	No	Yes	0.0000	No	No	0.0000	No	No	0.0005	0	1
NAGARCHstd	0.0014	No	Yes	0.0010	No	No	0.0006	No	No	0.0010	0	1
gjrGARCHged	0.0014	Yes	Yes	0.0011	No	No	0.0011	No	No	0.0012	1	1
gjrGARCHnig	0.0013	Yes	No	0.0010	No	No	0.0013	Yes	Yes	0.0012	2	1
gjrGARCHnorm	0.0015	Yes	No	0.0009	No	No	0.0010	No	No	0.0011	1	0
gjrGARCHstd	0.0000	No	No	0.0011	Yes	No	0.0012	Yes	No	0.0008	2	0
sGARCHged	0.0000	No	No	0.0000	No	No	0.0000	No	No	0.0000	0	0
sGARCHnig	0.0016	Yes	Yes	0.0014	Yes	Yes	0.0018	Yes	Yes	0.0016	3	3
sGARCHnorm	0.0021	Yes	Yes	0.0018	Yes	Yes	0.0023	Yes	Yes	0.0021	3	3
sGARCHstd	0.0018	Yes	Yes	0.0016	Yes	Yes	0.0020	Yes	Yes	0.0018	3	3
Grand Total	0.0115	8	9	0.0102	5	4	0.0007	5	4	0.0075	18	17

[1] ζ_j sig indicates whether the implied volatility parameter was significant at 10% level
[2] The LRT column gives an indication whether the null for the likelihood ratio test was rejected at 10% level
[3] The last 2 columns are count representation of the findings on the implied volatility parameter

In total, 60 models were used to evaluate the contribution of the three implied volatility measures. For the sample of models, 18 out of 60 specifications confirmed the significant influence of the external regressor parameter, while in 17 of these cases the LRT confirmed an improvement in fit. The measure which had the most influence in the fit of the different GARCH models was the implied volatility measure derived through the Black-Scholes model. The ζ_{IV} was significant in 8 out of 20 models while the other two measures had 5 out of 20 significant parameters. It also displayed a higher significance rate for the LRT which indicated that it helped with fit in 9 out of the 20 models. This is more than double than that of the SAVI and VIX.

For the standard GARCH specification, 3 out of 4 models confirms the contribution of the implied volatilties measure in estimating a better fit, both in terms of the significance of the $zeta_j$ parameter and LRT. The standard GARCH model with a GED error distribution specification however does not follow suite with its family of specification. It seems that when the fat-tailed characteristic of returns are accounted form, no suplementary information is added to aid the fit of the model.

In the case of the models that take into account leverage and assymetry effects, the influence of the implied volatility measures become less prominent. For the GJR-GARCH, in 6 of 12 models the external parameter ζ_j was significant, but of such a small magnitude, that it didnt improve the fit of the model as indicated by the LRT test where only in 2 of the 6 cases the fit improved. An interesting observation comes from the results of the NA-GARCH specification for the variance model. In 3 out of 12 cases the LRT indicated an improvement in fit, although the ζ_j parameter was insignificant. On closer inspection of this finding however it is revealed that the 10% cut-off value for significance excludes theses parameters in cross-tabulation, but in fact they are close to the cut-off value for significance[1]. The AVGARCH formulation indicated no relevency of any of the implied volatility measures for any of the 12 models.

The component GARCH model shows the biggest indications of influence from the the different implied volatility measures. The coefficient for the ζ_j parameter averaged 0.0449 and 0.09 for the significant models and indicated a very high influence in the case of the SAVI implied volatility measure with a coefficient of 0.1931.

8. Out-of sample forecast evalution

For out out of sample forecast consider the mechanism of forecasting using the basic GARCH model as formalized in (6) given the sample $t = 1, \ldots, T$. It can then be said that the best linear predictor of σ_{T+1} given information set $\mathcal{F}_t$ at time T is:

$$\begin{aligned} E[\sigma^2_{T+1}|\mathcal{F}_t] &= \alpha_0 + \alpha_1 E[\varepsilon^2_T|\mathcal{F}_t] + \beta_1 E[\sigma^2_T|\mathcal{F}_t] \\ &= \alpha_0 + \alpha_1 \varepsilon^2_t + \beta_1 \sigma^2_T \end{aligned} \tag{28}$$

[1]The p-values for the NAGARCH model with IV as external parameter with different error distributions are 0.1477 for norm, 0.1066 for std, 0.1631 for ged and 0.1151 for nig

Using the chain-rule of forecasting and $E[\varepsilon_T^2|\mathcal{F}_t] = E[\sigma_T^2|\mathcal{F}_t]$, eq (28) becomes:

$$\begin{aligned} E[\sigma_{T+2}^2|\mathcal{F}_t] &= \alpha_0 + (\alpha_1 + \beta_1)E[\sigma_{T+1}^2|\mathcal{F}_t] \\ &\vdots \\ E[\sigma_{T+k}^2|\mathcal{F}_t] &= \alpha_0 + (\alpha_1 + \beta_1)E[\sigma_{T+k-1}^2|\mathcal{F}_t] \\ &= \alpha_0 \sum_{i=0}^{k-1}(\alpha_1 + \beta_1)^i + (\alpha_1 + \beta_1)^{k-1}(\alpha_1\varepsilon_T^2 + \beta_1\sigma_T^2) \end{aligned} \tag{29}$$

In this paper an expanding window, or recursive method of forecasting is used with out-of-sample $N = 122$ ranging from 4 June 2013 to 20 Nov 2013. This uses an initial sample $t = 1, \ldots, T$ is used to estimate the model, producing a specified h-step ahead forecast out of sample. The sample is then increased by $T + 1$ and the model is re-estimated again producing h-step forecasts, but from $T + 1$, this continues till end of sample is reached and $\sigma^2 \in \{[1, \ldots, t + N, \ldots\ldots, t + h + N]\}$

Once the forecasting measure has been obtained, we assess their performance relative to a volatility benchmark. The forecast sample is then truncated to $N = 95$ in order to introduce a homogenous sample between 1-day, 5-day and 20-day horizon forecast for evaluations. Following Blair et. al. (2001) and Giot (2005) we use realized volatility, defined as the sum of squared daily returns:

$$RV_{k,t} = \sqrt{\sum_{j=1}^{k} r_{t+j}^2} \tag{30}$$

The performance of the various forecasters can now be determined by running the so called Mincer-Zarnowitz performance regression (Mincer and Zarnowitz, 1969) as well as the Diebold-Mariano (Diebold and Mariano, 1995) tests.

Figure (8) fig:r interval forecasts for the unrestricted model alongside the GARCHged forecasts when the three different implied volatility measures are included.

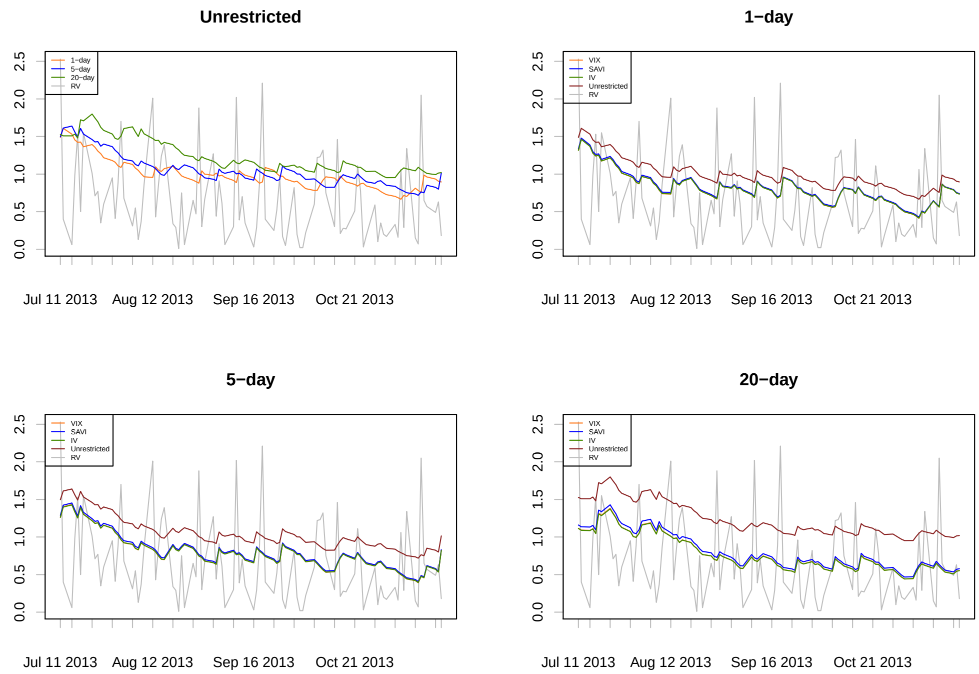

Figure 8: Forecast comparison

Although the inclusion of the implied volatility measure does seem to produce homogenous forecast for the 1-day and 5-day forecast, for the 20-day forecast, heterogeneity starts developing. Visually it seems that the VIX and IV seem to agree upon the implied volatility, while the SAVI measure forecasts marginally higher volatility. It is also noticable that the divergence between the unrestricted model and the general models have increased. This could indicate a bias towards the persistence parameter for the unrestricted model for the longer period forecasts.

8.1. Forecast Evaluation Statistic

In this section all permutations of specified GARCH with its different conditional distributions are tested against its unrestricted counterpart in an out-sample framework. By evaluating the forecast ability of the models with the set of implied volatility variables (SAVI, VIX and IV) as external regressors agains the unrestricted models, fit can be assessed and a statistical inference can be made whether the implied volatility measures do indeed contain information content which contributes towards volatility predictions. The hypothesis that is then tested is whether the market is pricing future volatility correctly, or from an asset management perspective, whether the market is efficient at pricing such assets as options.

8.1.1. Mincer-Zarnowitz

In a paper by Mincer and Zarnowitz (1969), they proposed methodology to evaluate forecasts was first formulated. The idea is to evaluate the forecast by approaching it in a regression fashion:

$$\hat{\sigma^2_{T+h}} = \alpha + \beta E_{i,T}[\sigma^2_{T+h}] + e_{i,T+h} \tag{31}$$

where .This method is often employed due to its simplicity in evaluation. High goodness of fit values R^2 give indication to accurate forecasts, while $\alpha = 0$ and $\beta = 1$ indicate unbiased forecasts. The modeld does however tend to underestimate the coefficient and so suffers from what is known as the attenuation bias resulting from an error-in-variable type problem that is created when estimated GARCH parameters are used to form $E_{i,T}[\sigma^2_{T+h}]$. We thus focus on the value for the goodness of fit measure R^2 as is often used for criteria in literature (Andersen et al., 2003). As the R^2 is being used as measure of evaluation it must be noted that, Andersen and Bollerslev (1998) found that as the quality for the proxy for independent variance term, $\sigma_{\hat{T+h}}$ deteriorates, the R^2 correlation metric punishes the evaluation.

Authors have raised concerns surrounding the effect of extreme observations biasing the forecast evalutaion due to strong leverage effect these observations have (Bollerslev and Wright, 2001). By transforming the variance[2], in actuallity the proxy for variance $\hat{\sigma}_{T+h}$ and $E_{i,T}[\sigma^2_{T+h}]$, a weakness emerges whereby the unbiased forecast for the latent unobserved variance are not generally unbiased for the transformed proxy measure, $\hat{\sigma}_{T+h}$. In this paper no transformation will be applied as results did not significant change due to trasnformation of the independent variable[3].

8.1.2. Diebold-Mariano

Using the recursive method for forecasting as laid out in eq (29), let $E_{i,T}[\sigma^2_{T+h}]$ denote the h-step ahead forecast of the volatility σ^2_{T+h} at time T

[2]Transformation suggestions include $log(\hat{\sigma}_{T+h})$ on $log(E_{i,T}[\sigma^2_{T+h}])$, or absolute returns $|r_t|$ on $\sqrt{E_{i,T}[\sigma^2_{T+h}]}$

[3]Results available on request from author

from a specific GARCH specification i. The corresponding forecast error is then denoted as $e_{i,T+h|T} = E_{i,T}[\sigma^2_{T+h}] - \sigma^2_{T+h}$. In this paper, the expected variance,$E_{i,T}[\sigma^2_{T+h}]$ is the realized volatility measure derived earlier. The most common forecast evaluation statistics in a N out-of sample foreast for the period $T = T+1, \ldots, T+N$ are:

$$MSE_i = \frac{1}{N} \sum_{j=T+1}^{T+N} e^2_{i,j+h|j}, \tag{32}$$

$$MAE_i = \frac{1}{N} \sum_{j=T+1}^{T+N} |e_{i,j+h|j}| \tag{33}$$

The intuition behind these measures dictate that the best model would produce the smallest forecast evaluation statistic, but these statistical values are random variables and a formal evaluation is prefered to infer model superiority. The forecast accuracy method is thus derived through a statistical approach derived by Diebold and Mariano (1995). It is a simple procedure to test the null of one model's forecast superiority over another by employing the traditional forecast evalution statistic.

Let $\{e_{1,j+h|j}\}^{T+N}_{T+1}$, and $\{e_{2,j+h|j}\}^{T+N}_{T+1}$ denote the errors in forecast generated by 2 different GARCH models. In our case, this represents the errors generated from the unrestricted mode without any external regressors and the errors from the model with the inclusion of the implied volatility measure. The forecast accuracy is measured through the use of a particular loss function. For the purpose of this paper, the squared error loss function $L(e_{i,T+h|T}) = (e_{i,T+h|T})^2$ is used. Other loss functions include the absolute error loss function $L(e_{i,T+h|T}) = |e_{i,T+h|T}|$. The Diebold-Mariano forecast evaluation test is based on the loss differential of the 2 models:

$$d_{T+h} = L(e_{1,T+h|T}) - L(e_2, T+h|T) \tag{34}$$

The DM statistic with the null of equal predictive accuracy, $H_0 : E[d_{T+h}] = 0$, is:

$$S = \frac{\bar{d}}{(\widehat{avar}(\bar{d})^{1/2}} \tag{35}$$

where $\bar{d} = N^{-1} \sum j = T+1^{T+N} d_{j+h}$, and $\widehat{avar}(\bar{d})$ is a consistent estimate of the asymptotic variance of $\sqrt{N}\bar{d}$. Due to the sample of loss differentials

$\{d_{j+h}\}_{T+1}^{T+N}$ being serially correlated for $h > 1$, Diebold and Mariano (1995) recommend the use of the Newey-West estimate for $\widehat{avar}(\bar{d})$ that corrects for the autocorrelation.

For S under the null hypothesis of equal predictive accuracy, has an asymptotic standard normal distribution which allows for the DM statistic to be used to test whether one model's forecast evaluation statistic is statistically different from another's (MSE_1 vs MSE_2).

9. Results

9.1. Mincer-Zarnowitz

Table (7) represents the results from eq (31). For evaluation, the regression was first applied using the forecasts from the unrestricted model and then compared to its respective counterpart with the implied volatility included. Comparison involved employing the R^2 and evaluating whether the model which included the implied volatility measure had a higher R^2 for the out-of-sample forecast period. In terms of overall performance, the SAVI measure produced the most cases whereby the goodness of fit was improved. This meant an improvement of 24 estimation fits on 60 estimations in total. The highest percentage of improvement was in the long term forecast of $t+20$ where 18 out of 20 forecasts' R^2 measure improved when the SAVI implied volatility index was used as an external regressor. This also reflected in the other two measure where the $t + 20$ forecast saw an improvement of 16 out of 20 estimations when the VIX was included and 15 out of 20 in the case of the IV measure derived from Black-Scholes model.

For the period $t + 5$ improvements where the least observed with only 2 cases of improvements being observed for all implied volatility measures. These improvements were also observed within one specification of the models, namely the long-memory model, component GARCH.

Table 7: Mincer-Zarnowitz results[a]

Is $R^2_{info} > R^2_{unres}$?	IV			SAVI			VIX			ALL		
	$t+1$	$t+5$	$t+20$	$t+1$	$t+5$	$t+20$	$t+1$	$t+5$	$t+20$	$t+1$	$t+5$	$t+20$
AVGARCHged	No	No	Yes	No	No	Yes	No	No	Yes	0	0	3
AVGARCHnig	No	No	Yes	No	No	Yes	Yes	No	No	1	0	2
AVGARCHnorm	No	No	Yes	No	No	Yes	No	No	Yes	0	0	3
AVGARCHstd	No	No	Yes	Yes	No	Yes	Yes	No	Yes	2	0	3
csGARCHged	No	No	No	No	No	Yes	Yes	Yes	No	1	1	1
csGARCHnig	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	3	3	0
csGARCHnorm	No	No	No	Yes	No	Yes	No	No	No	1	0	1
csGARCHstd	Yes	Yes	No	Yes	Yes	No	No	No	Yes	2	2	1
NAGARCHged	No	No	Yes	No	No	Yes	No	No	Yes	0	0	3
NAGARCHnig	No	No	Yes	No	No	Yes	No	No	Yes	0	0	3
NAGARCHnorm	No	No	Yes	No	No	Yes	No	No	Yes	0	0	3
NAGARCHstd	No	No	Yes	No	No	Yes	No	No	Yes	0	0	3
gjrGARCHged	No	No	Yes	No	No	Yes	No	No	Yes	0	0	3
gjrGARCHnig	No	No	Yes	No	No	Yes	No	No	Yes	0	0	3
gjrGARCHnorm	No	No	Yes	No	No	Yes	No	No	Yes	0	0	3
gjrGARCHstd	No	No	Yes	No	No	Yes	No	No	Yes	0	0	3
sGARCHged	No	No	Yes	No	No	Yes	No	No	Yes	0	0	3
sGARCHnig	No	No	Yes	No	No	Yes	No	No	Yes	0	0	3
sGARCHnorm	No	No	Yes	No	No	Yes	No	No	Yes	0	0	3
sGARCHstd	No	No	No	No	No	Yes	No	No	Yes	0	0	2
Grand Total	2	2	15	4	2	18	4	2	16	10	6	49

[a] Was the R^2 for the model with the information content greater that the unrestricted model's R^2

[1] $t+i$ represents the forecast horizon used for evaluation

[2] The last 2 columns are count representation of the findings on the comparison of the R^2 measure

Table 8: Mincer-Zarnowitz R^2 results[a]

	$1-day$		$5-day$		$20-day$		ALL	
	Unrestricted[1]	Implied Vol	Unrestricted	Implied Vol	Unrestricted	Implied Vol	Unrestricted	Implied Vol
AVGARCHged	0.0485	0.0480	0.0235	0.0223	0.0170	0.0176	0.0296	0.0293
AVGARCHnig	0.0485	0.0490	0.0239	0.0231	0.0170	0.0176	0.0298	0.0299
AVGARCHnorm	0.0485	0.0480	0.0235	0.0225	0.0170	0.0176	0.0296	0.0294
AVGARCHstd	0.0484	0.0460	0.0235	0.0204	0.0170	0.0175	0.0296	0.0280
csGARCHged	0.0062	0.0062	0.0196	0.0196	0.0406	0.0406	0.0221	0.0221
csGARCHnig	0.0073	0.0151	0.0208	0.0225	0.0401	0.0142	0.0227	0.0173
csGARCHnorm	0.0058	0.0058	0.0194	0.0194	0.0411	0.0411	0.0221	0.0221
csGARCHstd	0.0066	0.0112	0.0199	0.0210	0.0402	0.0228	0.0222	0.0183
NAGARCHged	0.0533	0.0479	0.0242	0.0222	0.0200	0.0223	0.0325	0.0308
NAGARCHnig	0.0560	0.0503	0.0255	0.0237	0.0196	0.0222	0.0337	0.0321
NAGARCHnorm	0.0531	0.0478	0.0242	0.0222	0.0201	0.0222	0.0324	0.0308
NAGARCHstd	0.0537	0.0472	0.0244	0.0220	0.0200	0.0227	0.0327	0.0307
gjrGARCHged	0.0456	0.0436	0.0266	0.0236	0.0262	0.0267	0.0328	0.0313
gjrGARCHnig	0.0458	0.0436	0.0270	0.0236	0.0265	0.0271	0.0331	0.0314
gjrGARCHnorm	0.0456	0.0438	0.0267	0.0239	0.0263	0.0267	0.0328	0.0315
gjrGARCHstd	0.0456	0.0431	0.0265	0.0228	0.0263	0.0269	0.0328	0.0309
sGARCHged	0.0106	0.0048	0.0115	0.0057	0.0335	0.0363	0.0185	0.0156
sGARCHnig	0.0112	0.0061	0.0124	0.0072	0.0339	0.0365	0.0192	0.0166
sGARCHnorm	0.0098	0.0040	0.0110	0.0050	0.0336	0.0365	0.0181	0.0152
sGARCHstd	0.0111	0.0073	0.0120	0.0082	0.0336	0.0354	0.0189	0.0170

[a] R^2 measures for the different forecast periods
[1] The unrestricted model represents the estimate for the standard model, while Implied Vol represents the model's forecast estimate with the inclusion of the implied volatility measure
[2] The last 2 columns are count representation of the findings on the comparison of the R^2 measure

Table (8) gives a better indication of the R^2 improvement by looking at the mean R^2 for a given period of forecast estimation for all models.

It is clear that from the low numbers, a linear approach does have serious drawbacks in the case where large outliers introduce low levels of correlation between the estimated volatility and the proxy realized volatility series. Despite the low goodness of fit numbers, an interesting phenomenon emerges in the R^2 numbers of the models that introduce non-linearity and assymetry into its specification. It is evident in the short term forecasts, $1-day$ and $5-day$, that the NAGARCH, AVGARCH and gjrGARCH produces better fit than the standard specification, be it the vanilla GARCH or long memory component GARCH. These models do however provide the best estimates in the long run $20-day$ forecast, with csGARCH and standard GARCH having R^2 estimates of 0.0411 and 0.0365 respectively.

This means of forecast evaluation has obvious weaknesses whereby the outliers can have a strong influence on the outcome of the R^2. It is thus imperative that a more statistical approach be employed.

9.2. Diebold-Mariano

In this section we employ the Diebold and Mariano (1995) test as a statical approach to acertain whether the implied volatility measure have any information content. We use the test statistic in eq (35) in a one-sided test whereby we test for the alternative hypothesis that the forecast with the implied volatility measure is the better forecast method. We are thus looking to reject the null hypothesis of equal forecast error to infer the validity in information content. In the statistical approach, the implied volatility measure have a much greater influence on forecast accuracy than observed with the Mincer-Zarnowitz approach. The main reason could be that the proxy for realized volatility, squared return values, are right-skewed distributed with very fat-tails, exaccerbating the leverage effect that biases linear models (See figure (A.9)). Table(9) presents the results for DM statistic > 1.96 as such, rejection of the null of equal forecast accuracy of the two models using the square root loss function[4].

The $20-day$ forecast evaluation with 50 out of 60 models showing improvement with the inclusion of the implied volatility measure is representative of what was observed in the Mincer-Zarnowitz approach with 49 out of 60 forecast's showing improvement. For the $1-day$ and $5-day$ forecasts, the same proportion of improvement was observed. Models show no improvement for their specific forecast periods were the AVGARCH and component GARCH models. For the component GARCH, the SAVI measure was the prominent implied volaltility measure showing improved forecasts for all periods and models except the $20-day$ forecast using a ged conditional distribution specification. The AVGARCH displayed the opposite trend with non-rejection for the null hypothesis for 9 out of the 12 evaluations.

[4] The absolute loss fucntion was also implemented, but results did not drastically differ. Although, for the component GARCH, non of the implied volatility measures improved forecast accuracy in the case of the absolute loss function specification.

Table 9: Diebold-Mariano results[a]

	IV			SAVI			VIX			ALL		
	$t+1$	$t+5$	$t+20$	$t+1$	$t+5$	$t+20$	$t+1$	$t+5$	$t+20$	$t+1$	$t+5$	$t+20$
AVGARCHstd	no	no	yes	yes	yes	yes	yes	yes	yes	2	2	3
AVGARCHnorm	yes	yes	yes	no	no	no	no	no	yes	1	1	2
AVGARCHnig	yes	yes	yes	no	no	no	yes	yes	yes	2	2	2
AVGARCHged	yes	yes	yes	no	no	no	no	no	yes	1	1	2
csGARCHstd	yes	yes	yes	yes	yes	yes	yes	no	no	3	2	2
csGARCHnorm	no	no	no	yes	yes	yes	no	no	no	1	1	1
csGARCHnig	yes	yes	no	yes	yes	yes	yes	yes	yes	3	3	2
csGARCHged	yes	yes	yes	yes	yes	no	no	no	no	2	2	1
NAGARCHstd	yes	yes	yes	yes	yes	yes	yes	yes	yes	3	3	3
NAGARCHnorm	yes	yes	yes	yes	yes	yes	yes	yes	yes	3	3	3
NAGARCHnig	yes	yes	yes	yes	yes	yes	yes	yes	yes	3	3	3
NAGARCHged	yes	yes	yes	yes	yes	yes	yes	yes	yes	3	3	3
gjrGARCHstd	yes	yes	yes	yes	yes	yes	yes	yes	yes	3	3	3
gjrGARCHnorm	yes	yes	yes	yes	yes	yes	yes	yes	yes	3	3	3
gjrGARCHnig	yes	yes	yes	yes	yes	yes	yes	yes	yes	3	3	3
gjrGARCHged	yes	yes	yes	yes	yes	yes	yes	yes	yes	3	3	3
sGARCHstd	no	no	no	yes	yes	yes	yes	yes	yes	2	2	2
sGARCHnorm	yes	yes	yes	yes	yes	yes	yes	yes	yes	3	3	3
sGARCHnig	yes	yes	yes	yes	yes	yes	yes	yes	yes	3	3	3
sGARCHged	yes	yes	yes	yes	yes	yes	yes	yes	yes	3	3	3
Grand Total	17	17	17	17	17	16	16	15	17	50	49	50

[a] Rejection of null hypothesis of equal forecast accuracy is labelled as "yes"

[1] $t+i$ represents the forecast horizon used for evaluation

[2] The last 2 columns are count representation on all findings

In terms of the weekly, $5 - day$ forecast, the VIX provided the least amount of information content to the model forecast by only improving on the unrestricted model in 15 out of 20 model forecasts compared to 17 out of 20 for the IV and VIX. The IV measure can be said to contain the most information content in this case due to the dispersion of non-rejection of the null accross models. There wasn't a sole model which rejectd the information content within its specification family, with the other models contradicting the findings, as was the case for the SAVI with the AVGARCH model and their various different specifications.

10. Conclusion

References

Akgiray, V. (1989). Conditional heteroscedasticity in time series of stock returns: Evidence and forecasts. *Journal of business*, 55–80.

Andersen, T. and T. Bollerslev (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review 39*(4), 885–905.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (1999). The distribution of stock return volatility. *Preprint*, 1–48.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2003, March). Modeling and Forecasting Realized Volatility. *Econometrica 71*(2), 579–625.

Baker, M. and J. Wurgler (2006). Investor sentiment and the cross-section of stock returns. *The Journal of Finance 61*(4), 1645–1680.

Becker, R. and A. E. Clements (2008). Are combination forecasts of s&p 500 volatility statistically superior? *International Journal of Forecasting 24*(1), 122–133.

Becker, R., A. E. Clements, and A. McClelland (2009). The jump component of s&p 500 volatility and the vix index. *Journal of Banking & Finance 33*(6), 1033–1038.

Becker, R., A. E. Clements, and S. I. White (2006). On the informational efficiency of s&p500 implied volatility. *The North American Journal of Economics and Finance 17*(2), 139–153.

Becker, R., A. E. Clements, and S. I. White (2007). Does implied volatility provide any information beyond that captured in model-based volatility forecasts? *Journal of Banking & Finance 31*(8), 2535–2549.

Beckers, S. (1981). Standard deviations implied in option prices as predictors of future stock price variability. *Journal of Banking & Finance 5*(3), 363–381.

Blair, B. J., S.-H. Poon, and S. J. Taylor (2010). Forecasting s&p 100 volatility: the incremental information content of implied volatilities and high-frequency index returns. In *Handbook of Quantitative Finance and Risk Management*, pp. 1333–1344. Springer.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics 31*(3), 307–327.

Bollerslev, T. (1987, August). A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return. *The Review of Economics and Statistics 69*(3), 542–47.

Bollerslev, T., R. Y. Chou, and K. F. Kroner (1992). Arch modeling in finance: A review of the theory and empirical evidence. *Journal of econometrics 52*(1), 5–59.

Bollerslev, T. and J. M. Wooldridge (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric reviews 11*(2), 143–172.

Bollerslev, T. and J. H. Wright (2001, November). High-Frequency Data, Frequency Domain Inference, And Volatility Forecasting. *The Review of Economics and Statistics 83*(4), 596–602.

Campbell, J. Y. and L. Hentschel (1992). No news is good news: An asymmetric model of changing volatility in stock returns. *Journal of financial Economics 31*(3), 281–318.

Canina, L. and S. Figlewski (1993). The informational content of implied volatility. *Review of Financial studies 6*(3), 659–681.

Chiras, D. P. and S. Manaster (1978). The information content of option prices and a test of market efficiency. *Journal of Financial Economics 6*(2), 213–234.

Christensen, B. J. and N. R. Prabhala (1998). The relation between implied and realized volatility. *Journal of Financial Economics 50*(2), 125–150.

Christie, A. A. (1982). The stochastic behavior of common stock variances: Value, leverage and interest rate effects. *Journal of financial Economics 10*(4), 407–432.

Day, T. E. and C. M. Lewis (1992). Stock market volatility and the information content of stock index options. *Journal of Econometrics 52*(1), 267–287.

Diebold, F. X. and R. S. Mariano (1995, July). Comparing Predictive Accuracy. *Journal of Business & Economic Statistics 13*(3), 253–63.

Engle, R. (2001). Garch 101: The use of arch/garch models in applied econometrics. *Journal of economic perspectives*, 157–168.

Engle, R. and V. K. Ng (1993a). Measuring and testing the impact of news on volatility. *Journal of Finance 48*(5), 1749–78.

Engle, R. F. (1990). Stock volatility and the crash of '87: Discussion. *The Review of Financial Studies 3*(1), pp. 103–106.

Engle, R. F. and G. G. J. Lee (1999). A Permanent and Transitory Component Model of Stock Return Volatility.

Engle, R. F. and V. K. Ng (1993b). Measuring and testing the impact of news on volatility. *The journal of finance 48*(5), 1749–1778.

Fama, E. F. (1965). The behaviour of stock market prices. *Journal of Business*, 34–105.

Fleming, J. (1998). The quality of market volatility forecasts implied by s&p 100 index option prices. *Journal of Empirical Finance 5*(4), 317–345.

Fleming, J., B. Ostdiek, and R. E. Whaley (1995). Predicting stock market volatility: a new measure. *Journal of Futures Markets 15*(3), 265–302.

French, K. R., G. W. Schwert, and R. F. Stambaugh (1987). Expected stock returns and volatility. *Journal of financial Economics 19*(1), 3–29.

Frijns, B., C. Tallau, and A. Tourani-Rad (2010). The information content of implied volatility: evidence from australia. *Journal of Futures Markets 30*(2), 134–155.

Ghalanos, A. (2014). *rugarch: Univariate GARCH models.* R package version 1.3-4.

Giot, P. (2002). The information content of implied volatility indexes for forecasting volatility and market risk. *Available at SSRN 362440*.

Glosten, L. R., R. Jagannathan, and D. E. Runkle (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.

Granger, C. W. and S.-H. Poon (2001). Forecasting financial market volatility: A review. *Available at SSRN 268866*.

Hentschel, L. (1995). All in the family nesting symmetric and asymmetric {GARCH} models. *Journal of Financial Economics 39*(1), 71 – 104.

Latane, H. A. and R. J. Rendleman (1976). Standard deviations of stock price ratios implied in option prices. *The Journal of Finance 31*(2), 369–381.

Mandelbrot, B. (1963). The variation of certain speculative prices. *The Journal of Business 36*.

Mayhew, S. and C. Stivers (2003). Stock return dynamics, option volume, and the information content of implied volatility. *Journal of Futures Markets 23*(7), 615–646.

Merton, R. C. (1980). On estimating the expected return on the market: An exploratory investigation. *Journal of financial economics 8*(4), 323–361.

Mincer, J. A. (1969). Economic forecasts and expectations: Analysis of forecasting behavior and performance. *NBER Books*.

Mincer, J. A. and V. Zarnowitz (1969, October). The Evaluation of Economic Forecasts. In *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance*, NBER Chapters, pp. 3–46. National Bureau of Economic Research, Inc.

Newey, W. K. and K. D. West (1986). A simple, positive semi-definite, heteroskedasticity and autocorrelationconsistent covariance matrix.

Poterba, J. M. and L. H. Summers (1984). The persistence of volatility and stock market fluctuations.

Samouilhan, N. and G. Shannon (2008). Forecasting volatility on the jse. *Investment Analysts Journal* (67), 19–28.

Taylor, S. (1986). *Modelling Financial Time Series*. John Wiley & Sons.

Taylor, S. J. (2007). Modelling financial time series.

Vlastakis, N. and R. N. Markellos (2012). Information demand and stock market volatility. *Journal of Banking & Finance 36*(6), 1808–1821.

Whaley, R. E. (1993). Derivatives on market volatility: Hedging tools long overdue. *The journal of Derivatives 1*(1), 71–84.

Whaley, R. E. (2000). The investor fear gauge. *The Journal of Portfolio Management 26*(3), 12–17.

Xekalaki, E. and S. Degiannakis (2010). *ARCH models for financial applications*. John Wiley & Sons.

Appendix A. Figures

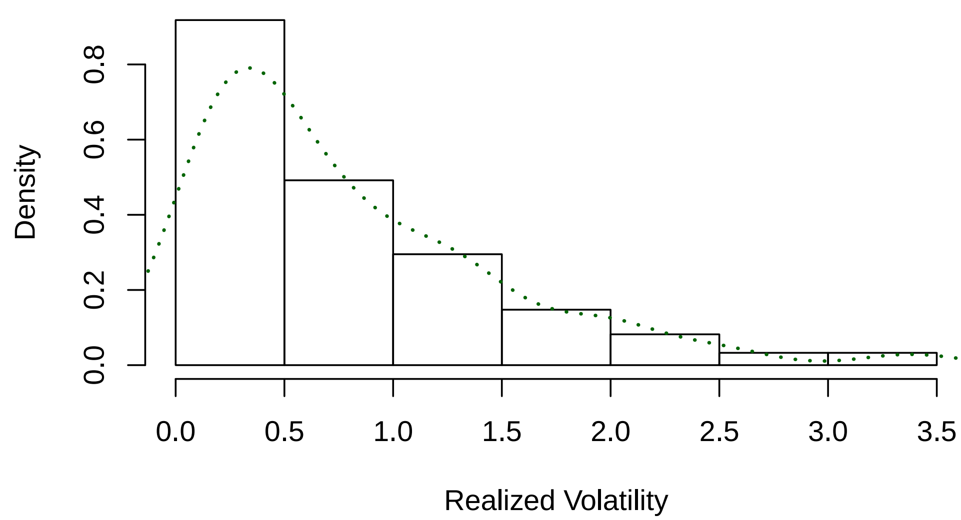

Figure A.9: Realized Volatility distribution