Erasmus University Rotterdam, Erasmus School of Economics

Econometrics and Operations Research

The information content of implied volatilities and MF volatility expectations: Evidence from options written on individual stocks

Kocak, K.

434504

Supervisor: Gong, M.

July 7, 2019

Abstract

We investigate different models for estimating the volatility of stock options for US firms and the S&P100 index: historical stock returns, MF volatility expectations and the ATM implied volatilities. If we look at the one-day-ahead forecasts, we observe that the ATM implied volatility performs the best more often, followed by the historical forecasts. For both sample periods are the option based forecasts better than historical volatility. Looking at the option based forecasts we see that the ATM implied volatility is better than the MF volatility expectations.

ERASMUS UNIVERSITEIT ROTTERDAM

Contents

1 Introduction

Measurements for index volatility have been important research topics for quite some time in financial literature. Most of the research shows, for the comparisons of the volatility forecasts that we get from historical stocks prices with forecasts obtained from option prices, that the accuracy of the volatility forecasts obtained from option prices is higher than the accuracy of volatility forecasts obtained from historical stock prices. This is, for example, shown in Christensen and Prabhala (1998) and Ederington and Guan (2002). From Andersen et al. (2001) we can see that option-based forecasts are frequently preferred over forecasts gained from the high-frequency realized volatility. The results for US stock indices are also shown in Jiang and Tian (2005), Blair et al. (2010) and Giot and Laurent (2007).

However, analogy of volatility forecasts for stock prices for individual firms are limited. Our contribution is an analogy of predictors based on historical stock prices and option-based forecasts for a large sample of US firms. Beside this, we will have a look at the analogy of model-free (MF) and at-the-money (ATM) predictors. The model-free forecasts cover the theoretical results of Demeterfi et al. (1999) and Britten-Jones and Neuberger (2000), these models apply combinations of option prices that do not rely on pricing formula. There is a shortcoming of application of the models on empirical data in many existing literature. The comparison of advantages and disadvantage of different models are rare.

This paper investigates how these different volatility measures, namely ATM implied volatility, historical daily returns and the MF volatility expectation, perform relative to each other. The performance we're going to compare will be estimating the volatility of stock returns. While doing so, we will reevaluate the information content of US index stock option prices, which the ATM implied volatility and the MF volatility expectation will use, relative to the historical daily returns. We will use two sample periods with a specific data selection criteria to apply these models to, one before the 2008 financial crisis and one after. This way we will investigate whether the financial crisis had any impact on the estimation of volatility of stock returns and at the same time test the robustness of the volatility measures.

After selecting our data in the aforementioned periods, we will explain the existing theoretical volatility measurements methodically. We will investigate these volatility measurements applied to selected data from two different sample periods, which will be our main contribution. This way we can reassess the earlier findings, particularly the consensus from studies that stock option prices are more informative than historical stock prices for index volatility measures. Our methodology concerning the sample periods will also allow us to investigate the significance of the influence made by the 2008 financial crisis when it comes to measuring index volatility, while also diversifying the conditions under which these models have to measure volatility. This way we will obtain more results to check and compare the relative performance of the aforementioned models.

We apply a selection criteria on our two data samples to improve data quality for applying the models to, which we will explain in detail in section 2. We choose for the period from January 2004 until December 2007 and January 2010 until December 2013. These sample periods are quite similar given the summary statistics in section 2. We have chosen these periods to investigate any differences of behaviour of the volatility measurements before and after the financial crisis.

The rest of the paper is organized as follows: section 2 lays down our selection criteria for the sample periods and illustrates the summary statistics. Section 3 covers the methodology we have used in our paper. We will discuss the three models for volatility measuring in this section. In this section we also introduce the regression methods for analyzing the information of various volatility estimates. Section 4 shows the results of our research: representing both the measurement models, comparisons between them and comparisons between the sample periods. We will conclude the work of this paper in section 5.

2 Data

We explain and summarize our selection procedure for our data set and describe the resulting data set using summary statistics.

We retrieved data from the OptionMetrics section of the Wharton Research Data Services institute. The sample period Taylor et al. (2010) uses is from January 1996 until December 1999. Our research is to investigate the sample period of January 2004 until December 2007, for futher investigation we will also look at the period of January 2010 until December 2013. We use, like in many other studies, the implied volatility that we obtain from the IvyDB database straight in our research. These implied volatilities are calculated based on the midpoint of the bid and ask prices in the order book. We account for non synchronous asset and option pricing by taking the average whenever a matching call and put implied volatility arise. We are not interested in options with less than a week to maturity for the sake of our analysis, because these option are close to leaving the market.

We use linear interpolation of the two closest zero-coupon LIBOR rates to obtain an interest rate corresponding to each option's expiration. The forward stock price $F_{O,T}$ with the same expiry date T as the options is then calculated from the difference of the future value of the current spot rate and the present value of all dividends until time T.

Firm selection undergoes a set of selection criteria. For starters, we're only interested in firms of whom options are traded throughout the whole sample period. Furthermore, the firm must have enough activity for us to be able to construct implied volatility curves for at least 98% of our sample data. To be able to estimate quadratic curves, at least three implied volatilities with three strike prices are required. Options who are the closest to their maturity are chosen, but if they do not meet the required amount of implied volatilities for quadratic curves we opt to chose the second closest option to its maturity date. When both the closest and the second closest option to its maturity date do not provide enough implied volatility data, we see that day as having missing data.

Table 1 presents summary statistics for 1-month estimates of the MF volatility expectation, ATM implied volatility, historical forecasts and the realized volatility. We first obtain the mean and standard deviation for the firms. Then the cross-sectional minimum, median and maximum values are calculated and displayed across the firms. We can see that the MF volatility expectation are higher on average compared to the ATM implied volatility.

2004-2007					2010-2013				
	Firms			S&P100		Firms			S&P100
	Min	Median	Max			Min	Median	Max	
Panel A: σ_{MF}					**Panel E**: σ_{MF}				
mean	0,24	0,48	0,81	0,23	mean	0,21	0,47	0,83	0,22
stdev	0,09	0,12	0,12	0,05	stdev	0,08	0,14	0,14	0,06
skewness	0,37	0,79	1,15	1,07	skewness	0,33	0,82	1,09	1,12
kurtosis	3,72	3,51	5,09	4,88	kurtosis	3,67	3,50	4,96	4,92
max	0,41	0,81	1,12	0,36	max	0,43	0,86	1,12	0,38
min	0,11	0,36	0,57	0,15	min	0,12	0,38	0,62	0,12
Panel B: σ_{ATM}					**Panel F**: σ_{ATM}				
mean	0,19	0,49	0,77	0,2	mean	0,22	0,53	0,82	0,17
stdev	0,06	0,09	0,11	0,04	stdev	0,05	0,11	0,14	0,05
skewness	0,29	0,71	1,15	0,75	skewness	0,28	0,62	1,06	0,68
kurtosis	3,46	3,31	4,68	3,62	kurtosis	3,41	3,13	4,49	3,43
max	0,41	0,71	1,09	0,34	max	0,49	0,73	1,21	0,36
min	0,12	0,34	0,52	0,12	min	0,11	0,32	0,56	0,14
Panel C: σ_{His}					**Panel G**: σ_{His}				
mean	0,16	0,52	0,72	0,15	mean	0,19	0,56	0,77	0,17
stdev	0,06	0,08	0,12	0,04	stdev	0,05	0,09	0,16	0,04
skewness	0,72	1,28	2,15	1,59	skewness	0,78	1,39	2,32	1,69
kurtosis	3,09	5,71	9,12	6,2	kurtosis	3,12	6,14	10.05	6,71
max	0,31	0,73	1,15	0,32	max	0,38	0,82	1,29	0,38
min	0,09	0,34	0,58	0,09	min	0,08	0,35	0,64	0,10
Panel D: σ_{RE}					**Panel H**: σ_{RE}				
mean	0,18	0,42	0,68	0,15	mean	0,23	0,48	0,72	0,19
stdev	0,08	0,11	0,15	0,05	stdev	0,10	0,15	0.17	0,03
skewness	0,52	1,03	1,68	1,95	skewness	0,65	1,22	1,76	1,82
kurtosis	2,98	4,21	6,62	8,09	kurtosis	2,42	3,89	6,42	7.98
max	0,34	0,86	1,18	0,36	max	0,39	0,92	1,32	0,42
min	0,10	0,22	0,39	0,09	min	0,13	0,29	0,46	0,13

Table 1: **Summary statistics for volatility measures for January 2004 to December 2007 and January 2010 to December 2013 .**These statistics are obtained from monthly observations of annualized volatility measures from the period 2004M01 to 2007M12 and 2010M01 to 2013M12. The denotations are as follows: estimates for the MF volatility expectation by σ_{MF}, the estimates for the ATM implied volatility by σ_{ATM}, the estimates from historical forecasts from the GARCH model by σ_{His} and the realized volatility calculated from daily high and low prices by σ_{RE}.

3 Methodology

3.1 Types of volatility forecasting

ARCH models are used for historical volatility forecasts, these models obtain the needed information from I_t. I_t is an information set that contains information of asset prices until time t. We have seen some examples of Blair et al. (2010). The conditional variance in the upcoming time period $t+1$, denoted as $h_t = var(r_{t+1}|I_t)$, is a forecast of the squared return in the next time period. Using this model has it advantages and disadvantages. Unfortunately, these forecasts rely on information of the past and not forward-looking. However, for estimating the parameters of the model we can use the maximum likelihood method. This method can also be used for the specification of h_t.

For deriving the risk-neutral expectation of the market are option based volatility essential. These volatilities contain also historical information. For the price of an underlying asset S_t follows: $dS = (r-q)Sdt + \sigma SdW$, with risk-free rate r, dividend yield q, the stochastic volatility σ_t and a time depending Wiener process W_t. The integrated squared volatility of the underlying asset from time 0 until the forecast horizon T is defined as $V_{0,T} = \int_0^T \sigma_t^2 dt$. This is equal to the quadratic var of the logarithm of the price process, because we assume here that there are no price jumps.
The theory of the MF volatility expectation is developed in several papers, like we mentioned in the introduction. At time 0 a set of European option prices is assumed to exist for an expiry time T.Britten-Jones and Neuberger (2000) show that the risk-neutral expectation of the integrated squared volatility is given by the following function:

$$E^Q\left[V_{0,T}\right] = 2e^{rT}\left[\int_0^{F_{0,T}} \frac{p(K,T)}{K^2}dK + \int_{F_{0,T}}^{\infty} \frac{c(K,T)}{K^2}dK\right], \tag{1}$$

with forward price $F_{0,T}$ at time 0 for a future transaction at time T, the expiry day. $c(K,T)$ and $p(K,T)$ are respectively the call and put prices at time T for a strike price K. The relation of equation above is as follows: right-hand side is equal to MF variance expectation, its square root is equivalent to the MF volatility expectation. To get the annualized values of these expressions we divide by T or $\sqrt{T}$. A continuous process is a major assumption for the MF volatility derivation given above.

The MF volatility expectations have some great advantages in comparison with Black-Scholes implied volatility. E.g. the MF volatility does not lean on any option pricing formula, which means that they do not need any assumptions of the dynamics of volatility. They also dodge leaning on a single strike price, this is not desirable due to the fact that implied volatilities differ a lot across strikes.

3.2 Specification of the models

The usage of the ARCH models brings many advantages. E.g. it gives an opportunity for acceptance of more observations and availability of maximum likelihood estimates to capture the parameters of the model.

However, despite some advantages we have to deal some disadvantages as well. If we look at the resting time to maturity, this is short compared with the time between multiple price observations. In order to maximize the unbiased of the volatility estimates obtained from the option prices, we look at both the one-day ahead forecast horizon and the time to maturity of the options horizon, like mentioned before. So in this paper two ARCH specifications will be evaluated. We evaluate our main ARCH specification for daily returns as follows:

$$r_t = \mu + \epsilon_t, \tag{2}$$

$$\epsilon_t = \sqrt{h_t} z_t, z_t \sim i.i.d.(0,1), \tag{3}$$

$$h_t = \frac{\omega + \alpha_1 \epsilon_{t-1}^2 + \alpha_2 S_{t-1} \epsilon_{t-1}^2}{1 - \beta L} + \frac{\gamma \sigma_{MF,t-1}^2}{1 - \beta_\gamma L} + \frac{\delta \sigma_{ATM,t-1}^2}{1 - \beta_\delta L} \tag{4}$$

Here h_t is the conditional variance of the returns in period t, L is the lag operator, s_{t-1} equals 1 if the error term of period $t-1$ is negative and equals zero otherwise. For the estimates of the MF volatility we use the term σ_{MF}; in the same way we use the term σ_{ATM} for the ATM implied volatility. These terms are obtained at time $t-1$ and divided by $\sqrt{252}$ for the annualized values.

We inspect different volatility models using different information sets. We first look at GJR-GARCH model, investigated in Glosten et al. (1993). This model only applies historical returns, hence we get $\gamma = \beta_\gamma = \delta = \beta_\delta = 0$. Secondly we look at the model obtained by information extracted from MF volatility expectations, hence $\alpha_1 = \alpha_2 = \beta = \delta = \beta_\delta = 0$. Finally, the model that is obtained by information presented by ATM implied volatility, hence $\alpha_1 = \alpha_2 = \beta = \gamma = \beta_\gamma = 0$. Characterized by assuming that:

$$z_t \sim N(\mu, \sigma).$$

We maximize the quasi-log-likelihood function to obtain the parameters. We also add some constraints on the parameters to make sure the conditional variance of the models remain positive:

$$\begin{aligned} \omega > 0; \\ \alpha_1 \geq 0, \alpha_1 + \alpha_2 \geq 0; \\ \beta \geq 0, \beta_\gamma \geq 0, \beta_{delta} \geq 0 \end{aligned}$$

The performance of the three models mentioned above depends on the value of their log-likelihood values. The conditional distributions of daily stock returns are better described by information with a higher log-likelihood value. When we rate specifications which have the same conditional mean, a higher rank indicates a more explanatory specification of the conditional variance and thus better performing forecasts of the realized variance.

We have seen in Blair et al. (2010) and Day and Lewis (1992) that comparable models with implied volatility are estimated into ARCH models for the S&P 100 index, as well as for individual stocks by Lamoureux and Lastrapes (1993). Previous studies presented for the individual US firms some asymmetric volatility effects, that is the reason for adopting the GJR model. We found this in studies of Cheung and Ng (1992) and Duffee (1995).

In a similar fashion to Canina and Figlewski (1993), Christensen and Prabhala (1998), Ederington and Guan (2002) and Jiang and Tian (2005) an encompassing and univariate regression for the realized volatility will be estimated for each individual firm. The first regression focuses on the relative effect of competing volatility estimates, while the latter can determine the explanatory content of each individual volatility estimation method.

The dependent variable in our analysis may turn out to be variance, volatility or, as shown in Jiang and Tian (2005), the logarithm of the latter. For the purpose of our research, each of these possible outcomes provide the same ranks for the forecasting instruments. Our preferences goes out to the logarithm of volatility, which is advantageous because of its approximate normal distribution characteristics which guarantee more reliable interpretations, as shown in Andersen et al. (2001).
For the regression we will use the following equation:

$$log(100\sigma_{RE,t,T}) = \beta_0+\beta_{His}log(100\sigma_{His,t,T})+\beta_{MF}log(100\sigma_{MF,t,T})+\beta_{ATM}log(100\sigma_{ATM,t,T})+\epsilon_{t,T} \quad (5)$$

$\sigma_{RE,t,T}$, defined by equation (6), is the realized volatility from time t to time T, it stands for the forecast quantity. The forecast of the historical volatility is obtained from the GJR-GARCH model and uses the information set up to time t, denoted by $\sigma_{Hist,t,T}$. For the MF volatility and ATM implied volatility we use respectively the terms $\sigma_{MF,t,T}$ and $\sigma_{ATM,t,T}$, of which the measures are non-overlapping. We use OLS estimates and take into account any existence of heteroscedasticity in the residual terms $\epsilon_{t,T}$: standard errors of White et al. (1980).

3.3 Forecasting tools

For the estimation of the ARCH models we use daily values, for both the MF volatility expectation and the ATM implied volatility. The ATM implied volatility equals to the strike prices that is available and nearest to the forward price. For calculating the MF volatility expectation, we extract a large amount of out-of-the-money option prices.

We use the formula of Parkinson (1980) to obtain the annualized realized volatility from day t until the option's maturity date T, for daily low and high stock prices such that:

$$\sqrt{\frac{252}{H}\sum_{j=2}^{H}\frac{log(h_{t+i})-log(l_{t+i})^2}{4log(2)}} \tag{6}$$

here h_t is the highest stock price at day t, l_t is the lowest stock price at day t and the number of days until expiry is denoted by H. Squaring the daily prices contributes to more precise measures of volatility than squared returns; consequently squared daily prices describe more fitting target quantities for comparisons of forecasts. Information set I_t at day t present the conditional variance h_{t+1} for day $t+1$, with a forecast of the variance until the day of expiry given by:

$$h_{t+1}+\sum_{j=2}^{H}E[h_{t+j|I_t}], \tag{7}$$

with forecast horizon H.

4 Results

Table 2 presents the correlations between one-month volatility forecasts: the MF forecast, the ATM forecast, the historical forecast of the volatility during the remaining lifetime of a set of option strikes and the realized volatility. The two volatility estimates, MF and ATM, are highly correlated as shown in Table 2. This correlation is less in the period from January 2010 to December 2013, however the two volatility estimates are still relative high. Across all the firms in the period from 2004 to 2007 the average correlation between the realized volatility and the MF expectation, the ATM implied volatility and the historical forecast, respectively, equals 0.523, 0.612 and 0.352. For the sample period from 2010 to 2013 the correlations decreased, namely 0.462, 0.423 and 0.296. The highest correlations are between the ATM and the MF volatilities, and their average value across the two periods equal 0.886.

Table 3 and 4 present the summary statistics of the sets of estimates and the S&P 100 index, from the ARCH specifications defined by (2-4). All the ARCH parameters are estimated using daily returns. Panel A summarizes estimates for the GJR-GARCH model; α_1 and α_2 are both measurements of the shocks ϵ_t. Here, α_2 only focuses on the negative

2004-2007

	σ_{MF} Firms				S&P100	σ_{ATM} Firms				S&P100
	Min	Median	Max	Mean		Min	Median	Max	Mean	
σ_{ATM}	0.921	0.919	0.932	0.941	0.976					
σ_{His}	0.535	0.512	0.523	0.568	0.899	0.572	0.562	0.559	0.579	0.893
σ_{RE}	0.512	0.492	0.503	0.523	0.656	0.476	0.452	0.460	0.612	
	σ_{His}									
σ_{RE}	0.356	0.348	0.375	0.352	0.466					

2010-2013

	σ_{MF} Firms				S&P100	σ_{ATM} Firms				S&P100
	Min	Median	Max	Mean		Min	Median	Max	Mean	
σ_{ATM}	0.811	0.789	0.712	0.831	0.952					
σ_{His}	0.475	0.467	0.470	0.498	0.845	0.498	0.506	0.486	0.512	0.872
σ_{RE}	0.462	0.482	0.459	0.462	0.601	0.401	0.412	0.386	0.423	0.586
	σ_{His}									
σ_{RE}	0.301	0.318	0.305	0.296	0.396					

Table 2: The correlations are between monthly observations of annualized measures of volatility. σ_{His}, σ_{MF}, σ_{ATM} and σ_{RE} are, respectively, the values of the historical, the MF volatility expectation and the ATM implied volatility from a GARCH model. All these volatility measures are for the remaining lifetimes of option contracts, which are approximately 1 month. The cross-sectional minimum, median, maximum and mean of each set of correlation statistics is calculated across the firms. The last column reports the correlations between monthly volatility observations from the index and another time series.

2004-2007	Mean	Std. dev.	Min	Median	Max	5%	S&P100
Panel A: GJR-GARCH model:							
μ x 10^3	0.96	0.91	0.24	0.87	1.66	11.4	0.91*
ω x 10^5	19.09	29.96	0.86	7.03	19.76	63.8	0.59*
α_1	0.04	0.09	0.00	0.04	0.09	20.1	0.00
α_2	0.10	0.21	0.02	0.09	0.16	43.0	0.21*
β	0.82	0.25	0.52	0.92	1.05	93.3	0.87*
L_{His}	2206	361	1896	2121	2476		3213
Panel B: MF expectations:							
μ x 10^3	0.78	0.93	0.28	0.75	1.46	11.4	0.76*
ω x 10^5	11.12	28.93	0.00	1.56	13.05	2.8	0.00*
γ	0.69	0.29	0.31	0.75	0.98	47.2	0.61
β_γ	0.24	0.23	0.00	0.06	0.45	9.3	0.00*
$\gamma/(1-\beta_\gamma)$	0.87	0.33	0.29	0.78	1.76		0.66
L_{MF}	2224	365	1882	2098	2438		3181
Panel C: ATM implied vol:							
μ x 10^3	0.82	0.96	0.40	0.76	1.42	13.6	0.92*
ω x 10^5	11.65	28.93	0.00	0.00	4.65	0.8	0.00*
δ	0.96	0.22	0.52	0.91	1.25	37.8	0.73
β_δ	0.16	0.23	0.00	0.10	0.25	8.1	0.00*
$\delta/(1-\beta_\delta)$	1.10	0.18	0.49	0.98	1.64		0.75
L_{ATM}	2228	362	1865	2210	2485		3168

Table 3: Daily stock returns σ_t by the ARCH specification: $\mathrm{r}_t = \mu + \epsilon_t, \epsilon_t = \sqrt{h_t} z_t, z_t \sim i.i.d.(0,1), h_t = \frac{\omega+\alpha_1\epsilon_{t-1}^2+\alpha_2 S_{t-1}\epsilon_{t-1}^2}{1-\beta L} + \frac{\gamma\sigma_{MF,t-1}^2}{1-\beta_\gamma L} + \frac{\delta\sigma_{ATM,t-1}^2}{1-\beta_\delta L}$.$\sigma_{MF}$ and σ_{ATM} are respectively the MF volatility expectation and the ATM implied volatility. The parameters are obtained by maximizing the log-likelihood function. Panel A, B and C presents respectively the GJR-GARCH model, MF expectation and the ATM implied volatility.

shocks, with s_t equals 1 if ϵ_t is negative and zero otherwise. Panel B provides results for the model which only uses the information contained in the time series of MF volatility. We see that in both periods the β_γ are close to zero or have a relatively small value. For γ we have values between 0.31 and 0.98 in the first period and for the second period this is between 0.28 and 0.86. Panel C summarizes estimates for the model which uses only the information contained in the ATM implied volatility series. The δ is relatively higher when compared to the γ of the MF volatility. However, the β_δ is also close to zero.

A higher log-likelihood value means a more accurate characterization of the distribution of the returns. We denote the maximum log-likelihoods of the models as follows: L_{His}, L_{MF} and L_{ATM}. All the values of the maximum log-likelihood are about similar. However, in both periods the maximum log-likelihood value for the ATM implied volatility has the highest value , followed by the MF expectation. The results and comparisons of

2010-2013	Mean	Std. dev.	Min	Median	Max	5%	S&P100
Panel A: GJR-GARCH model:							
μ x 10^3	0.93	0.90	0.22	0.91	1.52	9.8	0.86
ω x 10^5	18.02	31.96	1.52	9.46	19.76	59.4	0.53
α_1	0.05	0.09	0.00	0.04	0.09	18.6	0.00
α_2	0.09	0.21	0.00	0.08	0.14	41.6	0.19
β	0.87	0.27	0.55	0.98	1.10	91.2	0.84
L_{His}	1982	302	1612	1919	2166		2976
Panel B: MF expectations:							
μ x 10^3	0.83	0.96	0.43	0.84	1.62	10.2	0.69
ω x 10^5	10.33	25.31	0.05	1.69	14.95	2.18	0.00
γ	0.62	0.26	0.28	0.69	0.86	41.6	0.56
β_γ	0.24	0.23	0.00	0.06	0.45	9.13	0.00
$\gamma/(1-\beta_\gamma)$	0.87	0.33	0.29	0.78	1.76		0.66
L_{MF}	1986	322	1702	1823	2213		2890
Panel C: ATM implied vol:							
μ x 10^3	0.79	0.89	0.28	0.83	1.54	11.08	0.89*
ω x 10^5	13.51	26.32	0.00	0.00	14.65	0.96	0.00*
δ	0.79	0.19	0.45	0.86	1.04	42.1	0.68
β_δ	0.16	0.23	0.00	0.10	0.25	8.1	0.00*
$\delta/(1-\beta_\delta)$	1.10	0.18	0.49	0.98	1.64		0.75
L_{ATM}	2031	298	1724	1960	2126		2905

Table 4: Daily stock returns σ_t by the ARCH specification: $\mathrm{r}_t = \mu + \epsilon_t, \epsilon_t = \sqrt{h_t} z_t, z_t \sim i.i.d.(0,1), h_t = \frac{\omega+\alpha_1\epsilon^2_{t-1}+\alpha_2 S_{t-1}\epsilon^2_{t-1}}{1-\beta L} + \frac{\gamma\sigma^2_{MF,t-1}}{1-\beta_\gamma L} + \frac{\delta\sigma^2_{ATM,t-1}}{1-\beta_\delta L}$.$\sigma_{MF}$ and σ_{ATM} are respectively the MF volatility expectation and the ATM implied volatility. The parameters are obtained by maximizing the log-likelihood function. Panel A, B and C presents respectively the GJR-GARCH model, MF expectation and the ATM implied volatility.

	One-day-ahead forecasts		Options' life forecasts	
	2004-2007	2010-2013	2004-2007	2010-2013
Historical best	37.2%	36.1%	15.6%	17.3%
MF better than ATM	15.2%	14.2%	7.9%	10.1%
ATM better than MF	22.0%	21.9%	7.7%	7.2%
MF best	23.9%	26.4%	35.7%	37.4%
Historical better than ATM	19.1%	30.3%	32.4%	34.0%
ATM better than historical	4.8%	5.4%	3.3%	3.4%
ATM best	38.9%	37.5%	48.7%	45.3%
MF better than historical	32.2%	33.1%	42.2%	39.2%
Historical better than MF	6.7%	4.4%	6.5%	6.1%

Table 5: The percentages present how many of the firms satisfy the order of the best descriptions. We use the ARCH specifications for the one-day-ahead forecasts. The values given in the columns of the options' life forecasts are obtained from the univariate regressions.

the two periods are given in Table 5. For both periods is the ATM more often the best, for both one-day-ahead forecasts and option lifetime forecasts. Historical forecasts are slightly less than the ATM for one-day-ahead forecasts. However, for the option lifetime forecasts are the historical forecasts a lot worse.

Tables 6 and 7 presents the results of the regressions, given by equation 5, to analyze the realized volatility. It present the minimum, median, maximum and the S&P100, like before in this paper. Further, the table shows the R^2 and the sum of squared residuals. In both sample periods the ATM implied volatility has the highest R^2 value, followed by the MF volatility expectation. Seemingly the estimates from the option prices are better descriptive than historical stock returns.

2004-2007	β_0	β_{His}	β_{MF}	β_{ATM}	R^2	SSE
Panel A: His						
Mean	1.215	0.587			0.185	0.061
Min	0.191	0.327			0.049	0.047
Median	1.549	0.596			0.149	0.046
Max	2.609	0.929			0.258	0.065
S&P100	1.297	0.488			0.218	0.062
Panel B: MF						
Mean	0.891		0.617		0.237	0.056
Min	0.129		0.374		0.093	0.039
Median	0.693		0.614		0.219	0.049
Max	1.448		0.909		0.387	0.071
S&P100	0.120		0.749		0.498	0.051
Panel C: ATM						
Mean	0.761			0.714	0.228	0.061
Min	0.079			0.559	0.108	0.039
Median	0.704			0.761	0.205	0.049
Max	1.221			0.894	0.363	0.071
S&P100	0.203			0.914	0.536	0.070

Table 6: The most general estimated regression model for the logarithm of realized volatility is: $\ln(100\sigma_{RE,t,T} = \beta_0 + \beta_{His} + \ln(100\sigma_{His,t,T} + \beta_{MF}\ \ln(100\sigma_{MF,t,T}) + \beta_{ATM}\ \ln(100\sigma_{ATM,t,T}) + \epsilon_{t,T}$. With σ_{His}, σ_{MF}, σ_{ATM} and σ_{RE} , respectively, refer to the the historical forecast, the MF volatility expectation, the ATM option implied volatility and the realized volatility. The regressions are estimated by OLS for each firm. The minimum, median, maximum and mean are calculated for the different models.

2010-2013	β_0	β_{His}	β_{MF}	β_{ATM}	R^2	SSE
Panel A: His						
Mean	1.231	0.601			0.194	0.062
Min	0.191	0.327			0.056	0.048
Median	1.566	0.609			0.151	0.057
Max	2.616	0.934			0.266	0.071
S&P100	1.315	0.496			0.225	0.071
Panel B: MF						
Mean	0.901		0.624		0.247	0.059
Min	0.138		0.381		0.092	0.049
Median	0.701		0.621		0.231	0.059
Max	1.457		0.918		0.413	0.071
S&P100	0.125		0.757		0.512	0.051
Panel C: ATM						
Mean	0.771			0.729	0.238	0.065
Min	0.091			0.571	0.121	0.051
Median	0.721			0.771	0.223	0.061
Max	1.221			0.895	0.371	0.076
S&P100	0.221			0.931	0.523	0.071

Table 7: The most general estimated regression model for the logarithm of realized volatility is: $\ln(100\sigma_{RE,t,T} = \beta_0 + \beta_{His} + \ln(100\sigma_{His,t,T} + \beta_{MF}\ \ln(100\sigma_{MF,t,T}) + \beta_{ATM}\ \ln(100\sigma_{ATM,t,T}) + \epsilon_{t,T}$. With σ_{His}, σ_{MF}, σ_{ATM} and σ_{RE} , respectively, refer to the the historical forecast, the MF volatility expectation, the ATM option implied volatility and the realized volatility. The regressions are estimated by OLS for each firm. The minimum, median, maximum and mean are calculated for the different models.

5 Conclusion

In this paper we researched the information content of three volatility measures for two large sets of US firms. The ATM implied volatility and the, relatively new, MF volatility expectation make use of US index stock option prices, while the historical daily returns makes only use of the US stock index. Literature until now generally agrees that stock option prices are significantly more informational for measuring volatility of stock returns. While the results of our research show various advantages and disadvantages of both approaches, our main findings in the one-day-ahead forecasts contradict this statement.

If we look at the one-day-ahead forecasts, we observe that the historical forecasts performs comparably to the ATM implied volatility in both periods. When we extend the horizon to the maturity of the option however, both the MF volatility expectations and the ATM implied volatilties outperform the historical forecasts significantly. Looking within the option based forecasts, we see that the ATM implied volatility performs slightly better than the MF volatility expectations in both periods and for both horizons. Our paper shows that MF volatility expectation, a model which does not rely on any pricing formula, does not realize its theoretical potential.

By selecting our data with specific criteria, mainly concerning the high level of trade activity for an option, we established a high quality data set for the models using the stock option prices. In our results we found that stock options which had high trading activity enabled the models using stock option prices to more often outperform the historical forecasts. When the options had a relatively low trading activity however, the usage of historical returns yielded in better forecasts.

For the comparison of the two sample periods we look at the log-likelihood of the estimation, in general we see a lower value for the log-likelihood, this indicates a worse fit of the models than for the period before. The correlation of the volatility forecasts are decreased when we investigate the two sample periods. This may be attributable to the financial crisis, which made stock volatility less predictable. For more clear conclusions about the effect of the financial crisis, further research is required.

6 Bibliography

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Ebens, H. (2001). The distribution of realized stock return volatility. *Journal of financial economics*, 61(1):43–76.

Blair, B. J., Poon, S.-H., and Taylor, S. J. (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*, pages 1333–1344. Springer.

Britten-Jones, M. and Neuberger, A. (2000). Option prices, implied price processes, and stochastic volatility. *The Journal of Finance*, 55(2):839–866.

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

Cheung, Y.-W. and Ng, L. K. (1992). Stock price dynamics and firm size: An empirical investigation. *The Journal of Finance*, 47(5):1985–1997.

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

Day, T. E. and Lewis, C. M. (1992). Stock market volatility and the information content of stock index options. *Journal of Econometrics*, 52(1-2):267–287.

Demeterfi, K., Derman, E., Kamal, M., and Zou, J. (1999). A guide to volatility and variance swaps. *The Journal of Derivatives*, 6(4):9–32.

Duffee, G. R. (1995). Stock returns and volatility a firm-level analysis. *Journal of Financial Economics*, 37(3):399–420.

Ederington, L. H. and Guan, W. (2002). Measuring implied volatility: Is an average better? which average? *Journal of Futures Markets: Futures, Options, and Other Derivative Products*, 22(9):811–837.

Giot, P. and Laurent, S. (2007). The information content of implied volatility in light of the jump/continuous decomposition of realized volatility. *Journal of Futures Markets: Futures, Options, and Other Derivative Products*, 27(4):337–359.

Glosten, L. R., Jagannathan, R., and Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The journal of finance*, 48(5):1779–1801.

Jiang, G. J. and Tian, Y. S. (2005). The model-free implied volatility and its information content. *The Review of Financial Studies*, 18(4):1305–1342.

Lamoureux, C. G. and Lastrapes, W. D. (1993). Forecasting stock-return variance: Toward an understanding of stochastic implied volatilities. *The Review of Financial Studies*, 6(2):293–326.

Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *Journal of business*, pages 61–65.

Taylor, S. J., Yadav, P. K., and Zhang, Y. (2010). The information content of implied volatilities and model-free volatility expectations: Evidence from options written on individual stocks. *Journal of Banking & Finance*, 34(4):871–881.

White, H. et al. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *econometrica*, 48(4):817–838.