Original Article

Volatility forecasting and liquidity: Evidence from individual stocks

Received (in revised form): 22 May 2009

Peter Brous

is Professor of Finance at Seattle University. He earned a PhD in Finance from the University of Oregon. He has published articles in the *Journal of Finance*, the *Journal of Financial Economics*, the *Journal of Financial and Quantitative Analysis* and the *Journal of Accounting Research*. He has served as an expert witness in over 30 cases based on his expertise in valuing employee stock options and business valuation. His areas of research include corporate performance measures, capital budgeting, corporate financing decisions and real option analysis.

Ufuk Ince

is Visiting Assistant Professor of Finance at Seattle University. He received a PhD in Finance from Georgia State University. His area of research is corporate finance with a focus in financial risk management and derivatives use. He is the recipient of several National Science Foundation grants that examine financial modeling of seismic risk mitigation investments. He is a holder of the Chartered Financial Analyst designation and previously worked as a researcher at the Federal Reserve Bank of Atlanta.

Ivilina Popova

is Associate Professor of Finance at Texas State University – San Marcos. She received a PhD in Operations Research from Case Western Reserve University. Before academia, she spent several years working for Deutsche Asset Management in New York and Koch Capital Markets in Houston. She has published articles and conducts research in the areas of pricing derivatives, hedge fund portfolio construction and dynamic trading strategies.

Correspondence: Ivilina Popova, Department of Finance & Economics, McCoy College of Business Administration, Texas State University – San Marcos, San Marcos, Texas 78666, USA

ABSTRACT We evaluate the accuracy of four historical volatility measures and option-implied volatility in predicting future volatility using 10-year daily data for 92 stocks. In addition, we examine the effect of liquidity (proxied by option volume and stock market capitalization) on forecast accuracy. Our results suggest that forecast accuracy using implied volatility compared to historical measures is in general inferior, but for larger firms and higher option trading volume firms implied volatility forecasts are more accurate than historically based forecasts. Among the four historical measures, Parkinson's extreme value method and adjusted mean absolute difference perform best.

Journal of Derivatives & Hedge Funds (2010) 16, 144–159. doi:10.1057/jdhf.2009.17

Keywords: volatility forecasting; options; liquidity

INTRODUCTION

With the proliferation of financial options as an asset class, pricing them accurately has become an essential undertaking for market participants. Option traders rely on a variety of option pricing models and methods to quickly identify arbitrage opportunities. Corporate managers, who use options to hedge their firms' exposure to commodity, currency or interest rate risk, count on similar valuation techniques to implement their hedging strategies. Additionally, companies granting employee stock options need to expense and therefore value this form of compensation as it has been required by the US GAAP since 2005. Finally, when corporate investments are analyzed as real options, option valuation is necessary. In these option pricing applications, it is necessary to estimate future volatility for the underlying return series. Not only is the value of the option sensitive to the volatility input, but this input is also typically the most difficult to estimate.

There are two basic approaches in estimating future volatility. One approach uses the volatility of historical returns as a proxy for future volatility, whereas the other relies on volatility implied from option prices. Implied volatility estimates are generated using a pricing model that incorporates relevant past and current information about underlying return volatility. Therefore, one would expect implied volatility as a forward-looking measure to provide a better estimate of the future volatility than estimates using historical data.

There are a variety of volatility estimation techniques using historical stock price time series. In addition to the traditional squared difference method, there are measures based on absolute differences and extreme values, and numerous econometric models based on

non-constant volatility, such as ARCH, introduced by Engle.¹ The appeal of absolute return difference over squared return difference is the reduction in the influence of return outliers. Extreme value measures have the advantage of incorporating more information than measures based on closing prices because they also make use of intra-day high and low prices in estimating future volatility. Econometric approaches based on non-constant volatility, such as ARCH and GARCH models, capture the 'clustering' of volatility over time (see the study by Bollerslev *et al*² for a survey).

This study is focused on improving the understanding of which volatility forecasting method, historically based or implied, provides more accurate volatility forecasts. We contribute to the literature by examining the accuracy of four relatively simple historical measures of volatility, by examining a larger sample of individual stocks and by delving deeper into the effect of liquidity on the forecast accuracy of these methods. Our study provides thorough statistical analysis of the relative accuracy of volatility forecasts, and proposes specific practical guidelines for estimation of future volatility.

Focusing on individual securities rather than indexes allows us to compare the relative accuracy of alternative methods as a function of liquidity. Specifically, we examine whether two proxies for liquidity, option trading volume and market capitalization, affect the relative accuracy of volatility forecasts. Prior studies by Mayhew and Stivers³ have found that high option volume is related to improved implied volatility estimates. Market capitalization has not been previously examined as a factor in explaining the accuracy of volatility forecasts for individual stocks. We conjecture that larger firms will exhibit more efficient stock prices, and as a result

will have improved volatility estimates. This conjecture is supported by Chordia *et al.*,⁴ who provide evidence that the degree of market efficiency varies with firm size, and that the market for larger firms is more efficient.

Our results based on the full sample suggest that, on average, implied volatility forecasts are less accurate than all four historical measures. Additionally, among the four historical volatility metrics, we document that the extreme value method developed by Parkinson⁵ and the adjusted mean absolute deviation (AMAD) method provide the most accurate historically based volatility forecasts. Evidence regarding which historical volatility forecast is more accurate is particularly relevant when the firm does not have publicly traded options contracts, in which case implied volatility would not be a feasible method for forecasting volatility. When we split our sample based on proxies for liquidity (option volume and market capitalization), however, we find that implied volatility forecasts are more accurate than historical volatility estimates for larger firms and for firms with high option volume, and are less accurate for smaller firms and for firms with relatively low option volume.

LITERATURE REVIEW

The existing literature examining alternative measures of forecasting volatility is extensive. This review focuses on a set of papers comparing the accuracy or predictive powers of implied and historical measures of volatility. (For a more detailed literature review on the topic, see, for example, the study by Figlewski⁶ and Poon and Granger.⁷) The majority of studies analyze the implied volatility of stock indexes (S&P 100 and S&P 500), with only a few analyzing the implied volatility of individual stocks. These studies were

conducted with relatively small samples of individual stocks.

An early study by Canina and Figlewski,⁸ examining S&P 100 index options, finds that historical volatility contains relevant information regarding future variance that is not contained in the implied volatility estimate. A more recent study by Fleming,⁹ analyzing the implied volatility of the S&P 100 index, finds that implied volatility provides an upward biased forecast of actual future volatility. More relevant to our study and in contrast to previous research, he finds that the implied volatility dominates historical volatility (traditionally measured) in terms of forecasting future volatility. Christensen and Prabhala¹⁰ provide similar results, suggesting that implied volatilities estimated from the S&P 100 index options were stronger predictors of future volatilities than the standard difference of means historical volatility measure. They argue that their results are different from earlier findings by Canina and Figlewski⁸ because they examine longer time series and non-overlapping data. The survey article by Poon and Granger⁷ also concludes that option-implied volatility appears to have superior forecasting capability, outperforming many historical volatility models. Additionally, Corrado and Miller¹¹ examine the relationship between alternative measures of volatility and realized returns, and find that implied volatility of the S&P 500 index option has a positive and statistically significant relationship with realized returns, whereas historical measures of volatility are unrelated to realized returns. They measure historical volatility using a daily, high, low, close extreme value estimator.

A more recent study by Corrado and Truong¹² compares intra-day, high-low price extreme value estimators to implied volatility for several market indexes (S&P 100, S&P 500,

NASDAQ and Dow Jones Industrial Average). They find that neither the historical measure nor the implied volatility measure provides a better prediction of actual future volatilities. This result contradicts the preponderance of evidence in the literature suggesting that implied volatility forecasts are more accurate than historically based volatility forecasts for index options. We believe that Corrado and Truong's¹² unique finding is a result of the use of an extreme value volatility measure instead of the standard difference of means volatility measure.

Early studies on individual stock options' implied volatilities include those by Chiras and Manaster¹³ and Latane and Rendelman.¹⁴ They analyze a small group of stocks (fewer than 25) and conclude that implied volatility is more accurate than historically based volatility forecasts. Lamaoureux and Lastrapes¹⁵ examine options on 10 individual stocks that expired between 1982 and 1984 and find that historical volatility contains relevant information regarding future variance that is not contained in the implied volatility estimate. More recently, Mayhew and Stivers³ and Dennis *et al*¹⁶ analyze the information content of implied volatilities for 50 firms with the highest option volume on the CBOE between 1988 and 1995. They find that for firms with the highest option volume, implied volatility appears to perform better than historically based volatility forecasts. The evidence in their study suggests that high option volume leads to more efficient pricing of the options and thus more accurate implied volatility forecasts.

In a recent study, Ederington and Guan¹⁷ compare the predictive powers of two historical measures of volatility: the standard difference of means historical volatility measure

and the AMAD historical volatility measure. They find strong evidence that the AMAD is a more accurate predictor of future volatility than the standard difference of means measure. The relatively superior accuracy of the AMAD was found for future volatilities of both the S&P 500 index and five individual stocks.

In summary, for stock indexes, a majority of studies find that implied volatility is a better predictor of future volatility; however, Corrado and Truong,¹² examining a more sophisticated historical measure of volatility, find evidence inconsistent with previous research. Additionally, the scant evidence for individual stocks suggests that, on average, historical measures are more accurate than implied volatility forecasts; however, there is evidence that suggests that liquidity may have an effect on the relative accuracy of volatility forecasts. Overall, the literature has not provided a clear indication to practitioners as to which measure of volatility, historically based or implied, does a better job at forecasting future volatility.

METHODOLOGY

To conduct our tests, we generate volatility forecasts from four different measures of historical volatility and a volume-weighted measure of implied volatility. The volatility forecasts are then compared to a measure of realized volatility (RE). RE for n trading days is measured by using the following formula:

$$RE(n)_t = \sqrt{\frac{252}{n-1} \sum_{j=0}^{n-1} \left[r_{t-j} - \left(\frac{\sum_{j=0}^{n-1} r_{t-j}}{n} \right) \right]^2} \quad (1)$$

where n is 30 trading days, return at time t is $r_t = \ln \frac{C_t}{C_{t-1}}$ and C_t the closing stock price at time t .

To measure historical volatility, we first compute the historical standard deviation (SD), based on n trading days by using the following formula:

$$SD(n)_t = \sqrt{\frac{252}{n-1} \sum_{j=0}^{n-1} \left[r_{t-j} - \left(\frac{\sum_{j=0}^{n-1} r_{t-j}}{n} \right) \right]^2} \quad (2)$$

We also use the AMAD introduced by Ederington and Guan,¹⁷ defined as:

$$AMAD(n)_t = \sqrt{\frac{252\pi}{2} \frac{1}{n} \sum_{j=0}^{n-1} \left| r_{t-j} - \frac{\sum_{j=0}^{n-1} r_{t-j}}{n} \right|} \quad (3)$$

We then consider two extreme value estimators. Parkinson's⁵ extreme value estimator (PEVE) is defined as:

$$PEVE(n)_t = \sqrt{\frac{252}{n} \times \sum_{j=0}^{n-1} \frac{1}{4 \ln 2} \left(\ln \frac{H_{t-j}}{L_{t-j}} \right)^2} \quad (4)$$

where H_t denotes the daily high and L_t denotes the daily low price.

Garman and Klass's¹⁸ composite extreme value estimator (GKEV) as modified by Yang and Zhang¹⁹ is defined as:

$$GKEV(n)_t = \sqrt{\frac{252}{n} \times \left[\left(\ln \frac{O_t}{C_{t-1}} \right)^2 + \frac{1}{2} \left(\ln \frac{H_t}{L_t} \right)^2 - (2 \ln 2 - 1) \left(\ln \frac{C_t}{O_t} \right)^2 \right]} \quad (5)$$

where O_t denotes the daily open price.

We construct an implied volatility forecast that incorporates not only option prices, but

option volume as well. As a result, we have a liquidity-based implied volatility, which we call a Volume Weighted Implied Volatility (VWIV).²⁰ To compute the 30-day forward-looking implied volatility, we use a set of options with maturities T_1 and T_2 chosen such that $T_1 < 30 < T_2$. We call the options with maturity T_1 the near-term options and the options with maturity T_2 the next-term options. Our objective is to use the near- and next-term options to create an estimator for a constant maturity of 30 days to expiration. Suppose that we have the implied volatility estimators for the near- and next-term options, σ_{T_1} and σ_{T_2} , respectively. Our procedure for obtaining σ_{T_1} and σ_{T_2} is unique in that it not only combines the information embedded in the reported individual option's implied volatilities, but also incorporates liquidity information represented by the options' volume.

Suppose that on a given day t there are N^{T_i} strikes for options with maturity $T_i, i = 1, 2$. Denoted by $Volume_j^{T_i}, j = 1, K, N^{T_i}$ the number of option contracts (volume) traded on that day for a particular strike $K_j, j = 1, K, N^{T_i}$. We compute the volatility for the options with maturity T_i by volume-weighting the individual implied volatilities for puts and calls for all traded strikes:

$$\sigma_{T_i} = \sum_{j=1}^{N^{T_i}} \frac{Volume_j^{T_i} \times \sigma_j^{T_i}}{Volume}, \quad (6)$$

where $Volume^{T_i} = \sum_{j=1}^{N^{T_i}} Volume_j^{T_i}$ is the total number of puts and calls traded on day t for options with maturity T_i , and $\sigma_j^{T_i}$ is the reported implied volatility for an option (put or call) with strike K_j and maturity T_i . In formula (6), contracts with the highest volume have the largest impact on the computed implied volatility.

Finally, we linearly interpolate $\sigma_{T_1}^2$ and $\sigma_{T_2}^2$ to obtain the 30-day implied volatility estimator by using a formula borrowed from the Chicago Board Options Exchange Volatility Index (VIX) methodology:^{21,22}

$$\text{VWIV} = \sqrt{\left[T_1 \sigma_{T_1}^2 \frac{T_2 - 30}{T_2 - T_1} + T_2 \sigma_{T_2}^2 \frac{30 - T_1}{T_2 - T_1} \right]} \times \frac{1}{30} \quad (7)$$

Substituting equation (6) into (7) for the two maturities gives the implied volatility estimate that is used in the empirical sections that follow.

DATA DESCRIPTION

Our initial sample consists of members of the S&P 100 index. Introduced in 1983, the S&P 100 measures large-cap company performance. Often known by its ticker symbol OEX, it is comprised of 100 major blue chip companies across diverse industry groups. The primary criterion for index inclusion is the availability of individual stock options for each constituent. Our initial stock universe is comprised of all S&P 100 stocks at the beginning of 2007. The daily historical stock prices are collected from the Center for Research in Security Prices (CRSP) database. To collect daily stock option prices we use IVY DB OptionMetrics.²³ Only 94 of the initial set of 100 stocks had options traded back to 1996. Additionally, two of the companies experienced long periods of missing open price data on the CRSP database. As a result, our sample is reduced to 92 stocks. For each stock, we collect daily historical information from January 1996 to April 2006, yielding approximately 10 years of individual stock and options data.

For each stock, on a daily basis, the following data items are collected: OPTIONID, ISSUER, DATE, FLAG, EXERCISE, EXDATE, VOLUME, STRIKE and IMPL_VOL. OPTIONID is the individual option identifier, ISSUER is the name of the underlying issuer, FLAG is C if the option is a call and P if the option is a put, EXERCISE is A for American exercise, EXDATE is the maturity date of the option, VOLUME is the number of the traded contracts, STRIKE is the strike price of the option, and IMPL_VOL is the implied volatility of the option contract. The implied volatility, as reported in the OptionMetrics database, is calculated using American or European models where appropriate. All option calculations use historical LIBOR/Eurodollar rates for interest rate inputs, and incorporate discrete dividend payments.

To bracket the 30-day maturity, we use expiration dates between 8 and 91 trading days. On a particular day, the near-term option is the one that expires within 30 days and the next-term option is the one that expires in more than 30 days. For every given maturity, the implied volatility is computed by volume-weighting the reported implied volatilities of all traded call and put options of the same maturity. By following the procedure described in the previous section, the near- and next-term implied volatility estimates for σ_1 and σ_2 can be calculated. We calculate the implied volatility estimate for the 30-day maturity using equation (6).

To compute the historical volatility estimates, we use daily stock prices from the CRSP database. For a particular stock, daily close, open, high and low prices are used. We then use $n = 80$ and $n = 30$ -trading day windows to obtain values of the volatility estimates given by equations (1)–(4) (Figure 1). The results for

Figure 1: Volatility forecast timeline. This figure shows the trading day-based timeline for each of the 33 forecast intervals listed in Table 1. Five different methods are used to forecast the realized volatility of 92 S&P 100 stocks over a 10-year period. Realized volatility is the standard deviation of stock prices using data from days 0 to +30. Four of the five volatility forecast methods are historical, using data from days 80 to 0: SD (standard deviation), AMAD (adjusted mean absolute deviation), PEVE (Parkinson's⁵ extreme value estimator) and GKEV (Garman-Klass¹⁸ estimator of volatility). The remaining volatility forecast method is VWIV (implied volatility). To calculate VWIV, we use implied volatilities based on day-0 option prices. Contracts that expire before and after day 30 are selected and their implied volatilities are weighted by contract trading volume.

the two different window sizes are similar, and we report only the 80-day window results in the following empirical section.²⁴

In addition to the daily stock prices, we obtain daily market capitalization data from the CRSP database. On each date, we report the average daily market cap measured over the previous 80-day historical window.

EMPIRICAL RESULTS

We design a time sequence of 33 non-overlapping periods, with each period starting 80 trading days after the previous starting point. The 33 starting dates are listed in Table 1.

At each starting point, we compute the average daily market cap, the average daily option contracts volume, the four different historical volatility estimates, the 30-day implied volatility estimate and the 30-day realized volatility (RE30).

Panel A of Table 2 presents the descriptive statistics for the average daily market cap and average daily options contract volume. Panel B of Table 2 presents the descriptive statistics for the four different historical volatility estimates (SD, AMAD, PEVE and GKEV, all computed based on an 80-day historical window), 30-day implied volatility (VWIV), and 30-day realized volatility (RE30). The traditional squared difference historical volatility measure (SD) generates the highest mean and median volatility, whereas PEVE and the implied volatility generate the lowest volatility estimates.

Panel A of Table 3 provides the correlation coefficient for the average daily market cap and average daily options contract volume. The variables are positively correlated (0.67) and the coefficient is statistically significantly different from zero. Panel B of Table 3 provides the correlation coefficients for the four historical measures, the implied measure of volatility and the RE. All historical measures are highly correlated with each other, with the lowest correlation coefficient equaling 0.99. Not surprisingly, the implied volatility is not as strongly correlated with the historical measures, as evidenced by coefficients ranging from 0.87 to 0.89. Additionally, the RE is highly correlated with all historical measures (0.98), but not as highly correlated with the implied volatility measure (0.86).

The focus of this research is the forecast accuracy of the various historical measures of volatility and measures of implied volatility.

Table 1: Forecast intervals

<i>Interval</i>	<i>Date</i>	<i>Interval</i>	<i>Date</i>	<i>Interval</i>	<i>Date</i>
$t=1$	4 January 1996	$t=12$	30 June 1999	$t=23$	31 December 2002
$t=2$	29 April 1996	$t=13$	22 October 1999	$t=24$	28 April 2003
$t=3$	21 August 1996	$t=14$	16 February 2000	$t=25$	20 August 2003
$t=4$	13 December 1996	$t=15$	12 June 2000	$t=26$	12 December 2003
$t=5$	10 April 1997	$t=16$	4 October 2000	$t=27$	8 April 2004
$t=6$	4 August 1997	$t=17$	30 January 2001	$t=28$	4 August 2004
$t=7$	25 November 1997	$t=18$	24 May 2001	$t=29$	16 November 2004
$t=8$	24 March 1998	$t=19$	24 September 2001	$t=30$	23 March 2005
$t=9$	17 July 1998	$t=20$	17 January 2002	$t=31$	18 July 2005
$t=10$	9 November 1998	$t=21$	14 May 2002	$t=32$	8 November 2005
$t=11$	8 March 1999	$t=22$	6 September 2002	$t=33$	7 March 2006

This table shows 33 non-overlapping 80-day time intervals. For each of the 92 S&P 100 stocks, daily price data from January 1996 to April 2006 are collected. Volatility forecasts (SD, AMAD, PEVE, GKEV) using the previous 80 days and realized and implied volatilities for the subsequent 30 days are computed with reference to the dates indicated in the table.

Panel A of Table 4 presents the distribution statistics for root mean square forecast errors (RMSFE) for all of our volatility forecasts based on the entire sample.²⁵ All four of the historically based forecasts of volatility have mean errors smaller than the mean error based on implied volatility forecasts. Interestingly, the implied volatility forecast is the most accurate (best) forecast for 37 per cent of the firms in our sample, the highest percentage of all of our volatility forecasts based on historical data.

Panel B of Table 4 presents statistical significance results from all possible combinations of paired tests across all volatility forecasts for the entire sample. The differences in RMSFE are computed as:

$$\begin{aligned} \text{Difference RMSFE} = \\ & \text{RMSFE (column volatility measure)} \\ & - \text{RMSFE (row volatility measure)} \end{aligned}$$

For example, the difference reported in the top left-hand corner of Panel B (0.00668) equals:

$$\begin{aligned} \text{Difference RMSFE} = \\ & \text{RMSFE (SD)} - \text{RMSFE(AMAD)} \\ & = 0.12299 - 0.11631 = 0.00668, \end{aligned}$$

and suggests that the forecast error based on the traditional SD measure is greater than the forecast error based on AMAD.

In Panel B of Table 4 (as well as Tables 5 and 6), the results depict the forecasting performance of the method in a particular column in comparison to each method in the five rows. A negative figure in a particular cell indicates that the method in that column performs better than the method in the corresponding row. If a column has statistically significant negative figures in each of its rows, the method in that column is better than every other method. Conversely, positive and statistically significant

Table 2: Sample descriptive statistics

		Average market cap (in \$1000)			Average option contracts volume		
Panel A: Liquidity variables							
Mean		54 807 739				4088	
Median		32 221 103				1591	
St. dev.		60 968 645				6072	
Inter-quartile range		55 446 880				3815	
Range (min-max)		318 732 686				29 635	
Panel B: Volatility variables							
Mean	33.00	30.97	28.87	33.36	29.32	31.36	
Median	31.67	28.96	27.13	31.71	28.23	29.71	
St. dev.	8.88	8.32	7.42	8.77	7.79	8.30	
Inter-quartile range	9.39	8.35	7.76	8.92	7.19	9.50	
Range (min-max)	53	50	45	52	43	54	

This table shows the descriptive statistics for the variables used in the study. Inter-quartile range indicates the difference between variable values for the 25th and 75th percentile. Panel A shows the characteristics of liquidity variables. Panel B shows the characteristics of volatility variables SD (standard deviation), AMAD (adjusted mean absolute deviation), PEVE (Parkinson's⁵ extreme value estimator), GKEV (Garman-Klass¹⁸ estimator of volatility), VWIV (implied volatility) and RE30 (30-day realized volatility).

Table 3: Pair-wise correlations

		Average option volume			Average market cap	
Panel A: Between liquidity measures						
Average option volume		1.00				—
Average market cap		0.67 ***				1.00
Panel B: Between volatility measures						
Historical SD	1.00	—	—	—	—	—
AMAD	0.99 ***	1.00	—	—	—	—
PEVE	0.99 ***	0.99 ***	1.00	—	—	—
GKEV	0.99 ***	0.99 ***	0.99 ***	1.00	—	—
Implied volatility	0.89 ***	0.89 ***	0.87 ***	0.87 ***	1.00	—
Realized volatility	0.98 ***	0.98 ***	0.98 ***	0.98 ***	0.86 ***	1.00

This table reports the pair-wise correlations between the variables used in the study. Panel A shows the correlations among three liquidity measures: average option volume, average market capitalization and average stock volume. Panel B shows the correlations among six volatility measures: historical standard deviation (SD), adjusted mean absolute deviation (AMAD), Parkinson's⁵ extreme value estimator (PEVE), Garman-Klass¹⁸ estimator of volatility (GKEV), option implied volatility (VWIV) and realized volatility. The symbol *** indicates statistical significance at less than 1 per cent level.

Table 4: Forecast errors: full sample

	<i>SD</i>	<i>AMAD</i>	<i>PEVE</i>	<i>GKEV</i>	<i>VWIV</i>
<i>Panel A: RMSFE descriptive statistics</i>					
Mean	0.12299	0.11631	0.11686	0.12361	0.13173
Median	0.10911	0.10349	0.10483	0.11135	0.11744
St. dev.	0.05411	0.05061	0.05018	0.05044	0.05911
Inter-quartile range	0.05125	0.04435	0.04488	0.04447	0.06575
Range (min–max)	0.31	0.31	0.31	0.30	0.31
% Best	3%	27%	18%	14%	37%
<i>Panel B: RMSFE differences in means across five methods</i>					
SD	—	−0.00668***	−0.00613***	0.00062	0.00874**
AMAD	0.00668***	—	0.00056	0.00730***	0.01542***
PEVE	0.00613***	−0.00056	—	0.00675***	0.01486***
GKEV	−0.00062	−0.00730***	−0.00675***	—	0.00811*
VWIV	−0.00874**	−0.01542***	−0.01486***	−0.00811*	—

For the full sample of 92 stocks in the S&P 100 index, the root mean squared forecast error (RMSFE) is calculated using each of the five volatility forecast methods: standard deviation (SD), adjusted mean absolute deviation (AMAD), Parkinson's⁵ extreme value estimator (PEVE), Garman–Klass¹⁸ composite estimator of volatility (GKEV) and option implied volatility (VWIV). Panel A shows several descriptive statistics for RMSFE, where the lowest mean and median errors are shown in bold. Inter-quartile range indicates the difference between RMSFEs for the 25th and 75th percentile. % Best indicates the portion of stocks for which a particular volatility estimator garnered the lowest RMSFE. Panel B shows the magnitudes and statistical significance levels for the RMSFE differences between each of the five methods. The null hypothesis is $\text{Mean}(\text{RMSFE}_{\text{COLUMN}}) - \text{Mean}(\text{RMSFE}_{\text{ROW}}) = 0$. Therefore, negative values that are statistically significant indicate a better performance (that is, lower RMSFE) for the column method in comparison to the row method. The symbols *, ** and *** indicate statistical significance at less than the 10, 5 and 1 per cent levels, respectively, using a two-tail paired-sample *t* test.

figures indicate that the method in that column performs the worst among the five forecast methods.

The results reported in Panel B show that the difference between the RMSFE for all four of the historically based forecasts of volatility and the RMSFE for the implied volatility forecast is negative and statistically significant (see far right column under the VWIV heading of Panel B).

These results suggest that the RMSFE based on implied volatility forecasts are statistically significantly less accurate than all four historically based forecasts. The results presented in Panel B also indicate that the RMSFE based on AMAD and PEVE are statistically significantly less than the RMSFE based on SD and GKEV but are not statistically significantly different from each other. According to these results, historically

based volatility forecasts estimated using AMAD and PEVE are more accurate than historically based volatility forecasts estimated using SD or GKEV and the implied volatility forecast; however, there is no difference between forecasts estimated using AMAD or PEVE.²⁶

We expect firms with more actively traded options to have more accurate implied volatility forecasts than thinly traded options because the option prices used to derive these forecasts are more efficient. The results presented in Table 5 provide volatility forecast errors for both implied and historical measures controlling for option trading activity.

Panel A of Table 5 presents the mean RMSFE for the top and bottom quartiles based on option trading volume. The implied volatility forecast is the most accurate for the top quartile and the least accurate for the bottom quartile relative to historical measures. Furthermore, the implied volatility forecast is the most accurate for 74 per cent of the firms in the top quartile, the highest percentage of all forecasts, but is the most accurate for only 9 per cent of the bottom quartile firms. As expected, the implied volatility errors for the highest option volume quartile are lower than the bottom quartile; in contrast, the historical volatility errors are greater for the top quartile than the lower quartile.²⁷

Panel B of Table 5 presents results from paired tests designed in a similar way as the tests in Table 4, Panel B. Here, the tests are performed not on the whole sample of firms, but on two sub-samples – the top- and bottom-quartile firms ranked by option volume. Once again, the differences in RMSFE are computed as:

$$\begin{aligned} \text{Difference RMSFE} = & \\ & \text{MSFE (column volatility measure)} \\ & - \text{RMSFE (row volatility measure)} \end{aligned}$$

For the top quartile, the mean implied volatility forecast's error is statistically significantly smaller than all of the historically based forecasts. For the bottom quartile, however, the mean implied volatility forecast error is statistically significantly greater than all of the historically based forecasts. These results are consistent with our conjecture that for firms with greater option liquidity, the forecasts based on implied volatility are more accurate than forecasts based on historical stock returns, and that for firms with less liquidity, forecasts based on implied volatility are less accurate than all forecasts based on historical prices. A common finding for the top and bottom quartile is that the historically based forecasts measured using PEVE and AMAD are statistically significantly more accurate than the historical forecast made using SD.

We also expect that large-cap firms will have better historical and implied volatility forecasts than small cap firms based on the empirical evidence that larger firms exhibit more efficient stock prices. The results presented in Table 6 provide volatility forecasts for both implied and historical measures controlling for market capitalization.

The interpretation of Table 6 is similar to that of Table 5, but instead of ranking the sample based on option trading volume, the sample is ranked by market cap. Panel A of Table 6 reports that for the top quartile based on market cap, the mean RMSFE based on implied volatility forecast is the smallest compared to the mean RMSFE based on all four historical volatility forecasts, and that the percentage of firms for which the RMSFE based on implied volatility forecasts was the most accurate is 57 per cent. Conversely, as reported in Panel A of Table 6, for the bottom quartile, the mean RMSFE based on

Table 5: Forecast errors: top and bottom quartiles ranked by option volume

	<i>SD</i>	<i>AMAD</i>	<i>PEVE</i>	<i>GKEV</i>	<i>VWIV</i>
Panel A: Average errors					
<i>Average errors</i>					
Top quartile	0.13968	0.13347	0.13406	0.14234	0.12170
Bottom quartile	0.11098	0.10163	0.10112	0.10754	0.14362
<i>% best in group</i>					
Top quartile	0%	22%	0%	4%	74%
Bottom quartile	0%	26%	26%	39%	9%
Panel B: Average error differences across five methods					
<i>Top quartile</i>					
SD	—	−0.00621***	−0.00562***	0.00266	−0.01798***
AMAD	0.00621***	—	0.00059	0.00887	−0.01177**
PEVE	0.00562***	−0.00059	—	0.00828***	−0.01236***
GKEV	−0.00266	−0.00887	−0.00828***	—	−0.02064***
VWIV	0.01798***	0.01177**	0.01236***	0.02064***	—
<i>Bottom quartile</i>					
SD	—	−0.00935**	−0.00986**	−0.00344	0.03264***
AMAD	0.00935**	—	−0.00051	0.00591	0.04199***
PEVE	0.00986**	0.00051	—	0.00642	0.0425***
GKEV	0.00344	−0.00591	−0.00642	—	0.03608***
VWIV	−0.03264***	−0.04199***	−0.04250***	−0.03608***	—

For the top and bottom quartiles (23 stocks each) based on option trading volume of the full sample of 92 stocks in the S&P 100 index, the root mean squared forecast error (RMSFE) is calculated using each of the five volatility forecast methods: standard deviation (SD), adjusted mean absolute deviation (AMAD), Parkinson's⁵ extreme value estimator (PEVE), Garman-Klass¹⁸ composite estimator of volatility (GKEV) and option implied volatility (VWIV). Panel A shows average top- and bottom-quartile RMSFEs for each of the five volatility forecast methods, where the lowest mean and median errors are shown in bold. % best in group indicates the portion of stocks for which a particular volatility estimator garnered the lowest RMSFE. Panel B shows the magnitudes and statistical significance levels for the RMSFE differences between each of the five methods. The null hypothesis is $\text{Mean}(\text{RMSFE}_{\text{COLUMN}}) - \text{Mean}(\text{RMSFE}_{\text{ROW}}) = 0$. Therefore, negative values that are statistically significant indicate a better performance (that is, lower RMSFE) for the column method in comparison to the row method. The symbols *, **, and *** indicate statistical significance at less than the 10, 5, and 1 per cent levels, respectively, using a two-tail paired-sample *t* test.

Table 6: Forecast errors: top and bottom quartiles ranked by market capitalization

	<i>SD</i>	<i>AMAD</i>	<i>PEVE</i>	<i>GKEV</i>	<i>VWIV</i>
Panel A: Average errors					
<i>Average errors</i>					
Top quartile	0.11573	0.11016	0.11138	0.11949	0.10457
Bottom quartile	0.14376	0.12847	0.12831	0.13571	0.17830
<i>% best in group</i>					
Top quartile	9%	22%	4%	9%	57%
Bottom quartile	0%	22%	22%	43%	13%
Panel B: Average error differences across five methods					
<i>Top quartile</i>					
SD	—	−0.00557***	−0.00435**	0.00376*	−0.01116**
AMAD	0.00557***	—	0.00122	0.00933***	−0.00559
PEVE	0.00435**	−0.00122	—	0.00811***	−0.00681**
GKEV	−0.00376*	−0.00933***	−0.00811***	—	−0.01492***
VWIV	0.01116**	0.00559	0.00681**	0.01492***	—
<i>Bottom quartile</i>					
SD	—	−0.01529**	−0.01545**	−0.00805**	0.03454**
AMAD	0.01529**	—	−0.00016	0.00724	0.04983***
PEVE	0.01545**	0.00016	—	0.00740	0.04999***
GKEV	0.00805**	−0.00724	−0.00740	—	0.04259***
VWIV	−0.03454**	−0.04983***	−0.04999***	−0.04259***	—

For the top and bottom quartiles (23 stocks each) based on market capitalization of the full sample of 92 stocks in the S&P 100 index, the root mean squared forecast error (RMSFE) is calculated using each of the five volatility forecast methods: standard deviation (SD), adjusted mean absolute deviation (AMAD), Parkinson's⁵ extreme value estimator (PEVE), Garman-Klass¹⁸ composite estimator of volatility (GKEV), and option implied volatility (VWIV). Panel A shows average top- and bottom-quartile RMSFEs for each of the five volatility forecast methods, where the lowest mean and median errors are shown in bold. % best in group indicates the portion of stocks for which a particular volatility estimator garnered the lowest RMSFE. Panel B shows the magnitudes and statistical significance levels for the RMSFE differences between each of the five methods. The null hypothesis is $\text{Mean}(\text{RMSFE}_{\text{COLUMN}}) - \text{Mean}(\text{RMSFE}_{\text{ROW}}) = 0$. Therefore, negative values that are statistically significant indicate a better performance (that is, lower RMSFE) for the column method in comparison to the row method. The symbols *, ** and *** indicate statistical significance at less than the 10, 5 and 1 per cent levels, respectively, using a two-tail paired-sample *t* test.

implied volatility forecast is the largest compared to the mean RMSFE based on all four historical volatility forecasts. As expected, both the implied and historical volatility estimates are better for large-cap stocks.

Panel B of Table 6 shows that for the top quartile, the mean RMSFE based on implied volatility forecasts is statistically significantly smaller than the mean RMSFE based on all four historical volatility forecasts (except for AMAD). For the bottom quartile, the mean RMSFE based on implied volatility forecasts is statistically significantly greater than the mean RMSFE based on all four historical volatility forecasts. Once again, PEVE and AMAD are statistically significantly more accurate than the historical forecast made using SD.

CONCLUSION

In theory, implied volatility estimates derived from current option prices should be more accurate than historically based estimates of future volatility because they are forward looking. This study examines the accuracy of forecasts based on four historical measures of volatility and implied volatility forecasts for a large sample of individual stocks. One unique feature of our study is the focus on individual stocks, when most existing studies in this area have been conducted using stock indexes. The few studies that have tested the accuracy of volatility forecasts for individual stocks are based on relatively small samples. Therefore, our use of a relatively large sample of individual stocks provides more powerful tests of the accuracy of volatility forecasts. Furthermore, using individual stocks allows us to examine whether proxies for liquidity affect the relative accuracy of implied volatilities.

Our results for the full sample suggest that, on average, implied volatility, although theoretically superior, is less accurate than historically based estimates of future volatility. When we divide our sample based on proxies for liquidity, however, we find that implied volatility forecasts are more accurate than historically based forecasts for more liquid firms, but are less accurate for the sample of less liquid firms. Additionally, we show that historically based volatility forecasts estimated using Parkinson's extreme value method and adjusted mean absolute difference method are, in general, more accurate than the alternative methods using historical returns for both more liquid and less liquid firms.

Our main contribution to the literature is the notion that when selecting a method for estimating future volatility, one should take into account the type of firm that is being analyzed. More specifically, for more liquid firms, volatility forecasts are likely to be more accurate if based on implied volatility estimates. On the other hand, for less liquid firms, either Parkinson's extreme value method or the adjusted mean absolute difference method based on historical returns is likely to provide the most accurate forecast of future volatility.

REFERENCES AND NOTES

- 1 Engle, R.F. (1982) Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica* 50: 987–1008.
- 2 Bollerslev, T., Chou, R. and Kroner, K. (1992) ARCH modeling in finance: A review of the theory and empirical evidence. *Journal of Econometrics* 52: 5–59.
- 3 Mayhew, S. and Stivers, C. (2003) Stock return dynamics, option volume, and the information content of implied volatility. *The Journal of Futures Markets* 23(7): 615–646.

- 4 Chordia, T., Roll, R. and Subrahmanyam, A. (2005) Evidence on the speed of convergence to market efficiency. *Journal of Financial Economics* 76: 271–292.
- 5 Parkinson, M. (1980) The extreme value method for estimating the variance of the rate of return. *Journal of Business* 53(1): 61–65.
- 6 Figlewski, S. (2004) Forecasting Volatility. NYU Stern School of Business. Working Paper.
- 7 Poon, S.H. and Granger, C.W.J. (2003) Forecasting volatility in financial markets: A review. *Journal of Economic Literature* XLI: 478–539.
- 8 Canina, L. and Figlewski, S. (1993) The informational content of implied volatility. *The Review of Financial Studies* 6(3): 659–681.
- 9 Fleming, J. (1998) The quality of market volatility forecasts implied by S&P 100 index options. *Journal of Empirical Finance* 5(4): 317–345.
- 10 Christensen, B.J. and Prabhala, N.R. (1998) The relation between implied and realized volatility. *Journal of Financial Economics* 50: 125–150.
- 11 Corrado, C.J. and Miller, T.W. (2005) The forecast quality of CBOE implied volatility indexes. *The Journal of Futures Markets* 25(4): 339–373.
- 12 Corrado, C.J. and Truong, C. (2007) Forecasting stock index volatility: Comparing implied volatility and the intraday high-low price range. *The Journal of Financial Research* XXX (2): 201–215.
- 13 Chiras, D. and Manaster, S. (1978) The information content of option prices and a test of market efficiency. *Journal of Financial Economics* 6: 213–234.
- 14 Latane, H. and Rendleman, R. (1976) Standard deviations of stock price ratios implied in option prices. *Journal of Finance* 31: 369–381.
- 15 Lamoureux, C. and Lastrapes, W. (1993) Forecasting stock return variance: Toward an understanding of stochastic implied volatilities. *The Review of Financial Studies* 6(2): 293–326.
- 16 Dennis, P., Mayhew, S. and Stivers, C. (2006) Stock returns, implied volatility innovations, and the asymmetric volatility phenomenon. *Journal of Financial and Quantitative Analysis* 41(2): 381–406.
- 17 Ederington, L. and Guan, W. (2006) Measuring historical volatility. *Journal of Applied Finance* 16(1): 5–14.
- 18 Garman, M.B. and Klass, M.J. (1980) On the estimation of security price volatilities from historical data. *The Journal of Business* 53(1): 67–78.
- 19 Yang, D. and Zhang, Q. (2000) Drift-independent volatility estimation based on high, low, open, and close prices. *The Journal of Business* 73(3): 477–491.
- 20 In stock trading, a similar procedure is followed to create a trading benchmark using intraday stock prices: Volume Weighted Average Price (VWAP) is calculated by adding up the dollars traded for every transaction (price multiplied by number of shares traded) and then dividing by the total shares traded for the day.
- 21 Chicago Board Options Exchange. (2003) The CBOE volatility index – VIX, <http://www.cboe.com/micro/VIX/vixwhite.pdf>, accessed 2 September 2009.
- 22 For more details, see the CBOE VIX White Paper.²¹
- 23 As described in Wharton Research Data Services (WRDS): ‘Ivy DB OptionMetrics is a comprehensive source of historical price and implied volatility data for the US equity and index options markets. Encompassing more than six years of data, Ivy DB OptionMetrics contains historical prices of options and their associated underlying instruments, correctly calculated implied volatilities, and option sensitivities’.
- 24 Results with the window size $n=30$ are available on request.
- 25 RMSFE is defined as the square root of the average sum of squared errors between the realized volatility and any of the computed forecasts: SD, AMAD, PEVE, GKEV and VWIV.
- 26 We also measured forecasts errors based on mean absolute forecast errors (MAFE), and the results are similar to errors based on RMSFE.
- 27 By further examining the firms in the highest option volume quartile, the above result makes sense. The majority of these stocks have high volatility, are technology based and are not in a steady state. For such firms, one can expect that history will not be a strong predictor of the future.