Implied versus Stochastic Volatility: Evidence from a small Market*

Alfred Lehar
Department of
Business Administration
University of Vienna

Günter Strobl
Wharton

Martin Scheicher
Oesterreichische Nationalbank

July 26, 2000

*The authors would like to thank David P. Brown, Phillipe Jorion, Duane Seppi, Neal Stoughton, Timo Teräsvirta, Josef Zechner, seminar participants at the 1997 European Finance Association Meetings and at UC Irvine and the helpful comments of two anonymous referees. Thanks also to ÖTOB, the Vienna Stock Exchange, and Kontrollbank for providing the necessary data. A VSX Research Grant from the Vienna Stock Exchange and Support from Oesterreichische Nationalbank, Proj Nr. 6052 are gratefully acknowledged.

Implied versus Stochastic Volatility: Evidence from a small Stock Exchange

Abstract

This paper compares commonly used predictors for the volatility of stock returns. The techniques studied are moving averages of squared returns, Generalized Autoregressive Conditional Heteroscedasticity (GARCH), Stochastic Volatility (SV) and the Implied Volatility (IV). We perform this evaluation for the Viennese market, which has low liquidity compared to other exchanges in Europe, North America or Asia. For the purpose of model selection, we use a variety of econometric criteria. Our primary result is that the ranking of the models is sensitive to the criterion for measuring forecast performance. Therefore across the variety of models which we estimate no clear winner emerges. The implied volatility is found to contain information, which is absent in returns-based forecasts. We discuss possible explanations for these results. Based on our findings we suggest practical consequences for the purpose of derivatives valuation and risk management.

1 Introduction

Financial econometrics distinguishes two methods to predict the future variability of stock returns. The first approach uses option prices. Given an option pricing model we can infer from option prices a volatility measure which based on market data. This trade-based volatility can be interpreted as the market's expectation of the variability of the asset until the option expires. The alternative forecast method makes use of the time series of past returns. In contrast to the implied volatility, these predictors use a sample of daily returns and an econometric model to predict the future variance. Examples for this second approach are moving averages of squared historical returns, Generalized Autoregressive Conditional Heteroscedasticity (GARCH) or Stochastic Volatility (SV) [1].

Among practitioners, risk managers show particular interest in the performance of volatility models. The variances of risk factors are vital components of models which predict the amount of money a bank may lose on its trading activities over a certain time horizon. When implementing such a model, a risk manager needs to know which volatility specification gives the best out-of-sample performance. The current benchmark in this area is the Exponentially Weighted Moving Average (EWMA) which is used in the RiskMetrics^{TM} methodology. In addition to risk managers, traders also require accurate predictions of the future variability of price changes. Here an example is the trading of structured products and exotic options. In the case of exotic options, plain-vanilla derivatives are frequently used as hedge instruments. Valuation errors caused by inappropriate variance models influence the construction of the hedge and therefore lead to incorrect portfolio compositions.

Given this importance of volatility forecasting among practitioners many authors have compared alternative models. But to date, the studies mainly focused

[1] See Palm (1996) for a survey on volatility models

on large markets like the US stock market or the Foreign Exchange market.[2] Based on this state of the literature our paper offers two contributions: We extend the available evidence by studying a small and relatively illiquid market, namely the Vienna Futures and Derivatives Exchange[3]. Existing research for smaller markets is very limited. However also risk managers in smaller markets need evidence on the fit of variance predictors. One of the few published studies on this topic is Adjaoute et al. (1999) on the Swiss equity options market. They find that the implied volatility outperforms a GARCH-based predictor. The second contribution of our paper is that we analyse a representative set of variance models. In addition to GARCH, moving average measures and implied volatilities we include stochastic volatility. Relative to previous work, this is an important methodological extension. So far, the available studies compared IVs only to GARCH and moving averages. However, a number of authors document the good fit of the SV model, cf. Kim et al. (1996) or Mahieu and Schotman (1998). SV is also of particular interest because its continuous-time specification is the returns-generating process for the option pricing model proposed by Hull and White (1987).

The primary result of our study concerns the ranking of the models. We find that it depends on the criterion for measuring forecast performance. Therefore across the variety of models which we analyse there is no unambiguous winner. All predictors show systematic forecast errors. Comparing simple and complex volatility models, we do not find a pronounced advantage of the more elabortae returns-based approaches. Regarding the importance of the trade-based volatility we document that it contains information, which is absent in returns-based

[2]See Bates (1996) or Figlewski (1997) for surveys. Some important studies are Harvey and Whaley (1991), Canina and Figlewski (1993), Lamoureux and Lastrapes (1993), Jorion (1995), Heynen and Kat (1994), Christensen and Prabhala (1998) or Fleming (1998).

[3]See Scheicher (1999) for a comparison of the in-sample fit of time series models for the main index of the Vienna Stock Exchange.

Table 1
Turnover for stock options (excluding index options) in number of contracts and in billions of USD as of 1998 for major world derivatives exchanges.

	Stock Option Volume (no of contracts)	Stock Option Notional Value (million USD)
CBOE (USA)	138,507,143	660,904.3
EUREX Zürich (Switzerland)	30,104,990	246,926.1
BOVESPA (Brazil)	39,673,509	234,206.0
PSE (USA)	58,930,521	211,099.5
EUREX Deutschland (Germany)	30,853,777	163,026.2
PHLX (USA)	37,473,343	153,754.8
ASXD (Australia)	8,072,691	47,666.5
IDEM (Italy)	1,307,436	33,671.1
Osaka (Japan)	363,901	20,197.8
HKSE (Hong Kong)	1,637,447	11,020.6
Toronto SE (Canada)	3,460,585	9,669.5
ÖTOB (Austria)	1,201,197	9,341.7
MEFF Variable (Spain)	2,695,206	7,709.3
BVRJ (Brazil)	6,367,570	6,751.3
BELFOX (Belgium)	364,885	4,040.1
Montreal SE (Canada)	1,374,961	2,427.4
Oslo (Norway)	883,045	1,827.9
FUTOP (Denmark)	4,622	47.7

source: International Federation of Stock Exchanges

forecasts. Hence for our sample the best method to forecast future variance is to combine both approaches. So the optimal predictor includes information from historical returns and from option prices.

The rest of this paper proceeds as follows: Section 2 describes our sample and Section 3 outlines our methods for computing the IV and the alternative volatility models. Section 4 details the results while Section 5 summarises the principal findings.

Table 2
Market characteristics of major US and European stock exchanges.

Market	Market Capitalisation	Turnover	Trading Costs (basis points)
NYSE	10271.90	7317.9 T	24.57
Nasdaq	2527.97	5518.9 R	30.64
Great Britain	2372.74	2888.0 R	51.88
Germany	1093.96	1491.8 R	29.70
France	991.48	587.9 T	27.63
Italy	569.73	486.5 T	29.84
Netherlands	603.18	409.5 R	34.56
Spain	402.16	640.3 R	24.57
Sweden	278.71	230.0 R	32.26
Belgium	245.66	60.9 T	33.21
Austria	35.78	18.7 T	51.29

Market capitalisation and turnover are in billions of USD as of 1998 (source: International Federation of Stock Exchanges). The turnover figures should be interpreted with caution due to different calculation methods. Turnover figures with a T show only transactions that pass through a trading system of an exchange, figures with a R include on- and off-market transactions. Trading costs is measured in basis points as of the 3^{rd} quarter of 1998 and is computed as the average sum of commissions, fees and market impact based on trade data on all global trades done by 135 institutional investors (source: Elkins/McSherry Co.,Inc.) The last column of this Table is taken from Pagano et al. (2000), Table 3.

2 Sample

The Vienna Options and Derivatives Exchange is quote driven, with a fully automatic screen-based trading system and ten market makers. It opened in 1991 and currently lists options on 13 stocks, a domestic and five eastern European indices and on a bond futures contract. On the derivatives exchange the average open interest in 1998 amounted to 165,672 contracts with a yearly turnover of 2.9 million contracts for all derivatives. The market capitalisation of the stock exchange was USD 35.78 bn in December 1998. It had 119 companies listed.

To set the Vienna market into perspective, Table 1 and Table 2 show key market characteristics of a variety of derivatives exchanges. Table 1 indicates that the majority of European exchanges which offer trading in equity derivatives lack the turnover of the large US markets. Prominent examples are Belgium (BELFOX), Spain (MEFF Renta Variable) or Denmark (FUTOP). A lot of European stock markets also show not only lower capitalisation and turnover as their US-counterparts but also higher trading costs, as it can been in Table 2. These transaction costs are of immanent importance for hedging derivative positions in the spot market. These observations question the general applicability of the available empirical methods to samples from smaller markets as they are common in Europe. In particular, as we will discuss, the estimation of the implied volatility needs special attention.

Our sample comprises transaction data from options on the following stocks: Creditanstalt Bankverein (CA), Austria's largest bank during the sample period, Energieversorgung Niederösterreich (EVN), a regional utility company, Oesterreichische Mineralölverwaltung (OMV), Austria's principal oil company, Verbund (VER), the principal utility company, and Wienerberger (WIE), Europe's largest brick producer. These are the most actively traded equity options on ÖTOB. The focus of our study is only on stocks because the index strongly suffers from the problem of infrequent trading (cf. Harvey and Whaley (1991)). Our high frequency observations start at January 2^{nd} 1995 and end at June 21^{st} 1996. The actual trading days differ for the individual stocks, because some were suspended from trading around important announcements. Trading hours on the derivatives market are from 9:00 a.m. to 3:00 p.m. The exchange usually offers options with six different strike prices and maturities of the next three months and the last month of the following quarter. Table 3 gives descriptive statistics for the option trades. It shows the number of trading days, the total number of observed trades and the minimum and maximum number of trades observed on a day. The last column also gives the number of different contracts in the sample.

On average there are 65.2 option trades per day in our sample, 99.8 for the most active stock (CA) and only 28.7 for the least liquid one (WIE). The maximum number of trades observed on a single day was 688 (VER) but there were also 4 days in our sample where there were no options traded at all for Wienerberger, EVN or Verbund. On 46 days less than 5 trades could be observed for specific options. Most trades occur at the money.

Additionally we have transactions data on trades in the underlying stocks from the Vienna Stock Exchange, described also in Table 3. On average there were 20.6 trades per day, the maximum number of trades per day was 56 (OMV) and there was only one day with no trade in a stock (also OMV). CA was the most actively traded stock on the exchange.

To compute implied volatilities we have to match option and stock trades in order to get all required parameters for inverting the option pricing formula. For each option trade, the most recent stock price was chosen, and those trades not satisfying the following criteria below, were filtered out:

1. The trade in the stock was not more than one hour before the trade in the option
2. The option has a maturity of more than seven days
3. The price, at which the option was traded, does not violate an arbitrage bound

This filter procedure reduces the sample-size from 118,362 to 86,864, consisting of 59,320 calls and 27,544 puts. The main reasons for the 27% drop are different trading hours on the stock market and on the options exchange and infrequent trading, causing the stock prices matching the option trade to be too old. Table 4 gives a summary of the final sample. Apart from the problem of infrequent trading, turnover is generally low. We can also observe that small trades are most frequent[4]

Dividends and adjustments to stock prices and strikes due to capital measures are taken into account when calculating implied volatilities. The riskless inter-

[4]e.g. for CA 50% of the transactions (measured in premia) are below USD 2000.

Table 3
Stock and options trades in the sample

underlying stock	trading days in sample	Trades total	Trades min/day	Trades max/day	Contracts	Days with less than five trades
CA stock	360	7832	5	52		0
CA options	360	35,926	11	423	254	0
EVN stocks	366	7518	5	45		0
EVN options	366	20,911	0	193	287	7
OMV stocks	364	9312	0	56		2
OMV options	364	32,443	4	366	300	1
VER stocks	364	6786	4	45		1
VER options	364	18,678	0	688	280	14
WIE stocks	362	6001	3	44		5
WIE options	362	10,404	0	148	259	24

Table 4
Sample size before and after filtering.

stock	observed option trades	sample after matching total	calls	puts
CA	35,926	25,591	18,261	7,330
EVN	20,911	15,561	10,290	5,271
OMV	32,443	24,554	17,151	7,403
VER	18,678	13,340	8,290	5,050
WIE	10,404	7,818	5,328	2,490
total	118,362	86,864	59,320	27,544

est rate is taken from daily Vienna Inter-Bank Offered Rate (VIBOR) quotes. Interest rates for overnight and one, three, six and twelve months were included.

In addition to the intraday sample described above we used stocks' closing prices from Jan 2^{nd} 1992 up to June 21^{st} to estimate the parameters of the time series based volatility models.

3 Methodology

3.1 Computing the Daily Implied Standard Deviation

In the literature the standard methodology for computing the IV, taken e.g. by Jorion (1995), is to equate the Black and Scholes (1973) theoretical price with the observed price for at the money options with around two weeks to maturity. This expression is then numerically solved for the single remaining unknown, namely volatility. The assumptions in this procedure are that the cash and derivatives markets comply with the efficient market hypothesis (EMH) and that the Black-Scholes model is correct. Therefore any results from studying the IVs must be interpreted under the two caveats of the "joint hypothesis" problem: EMH and validity of the Black-Scholes option pricing model. The first issue influences mainly the interpretation of the results but the second issue is of much greater concern for the performance of the method. We justify the use of the Black-Scholes model in a number of ways. In choosing the BS approach we follow the market convention because this method is the most frequent "rule of thumb strategy" (Canina and Figlewski (1993)) for estimating the expected volatility in many markets. Even if the true process is heteroscedastic, the variance implied by the BS model serves as an approximation to the unobserved "true" variance. For the case of time-varying volatility there are a number of competing models, e.g. Hull and White (1987), Wiggins (1987), Heston (1993) or Duan (1995). However among the alternative approaches there is no dominant method. A drawback is that the specification and estimation of these stochastic processes adds measurement errors which are absent in the BS model bevcause the number of parameters rises. Already within the BS model a large number of measurement errors are documented, c.f. Christensen and Prabhala (1998). There is a pronounced bid/ask bounce and we observe non-synchronous trading, i.e. the intervals between a trade on the derivatives exchange and on the cash market are sometimes quite long. Other problems are price discreteness and the violation of the no-arbitrage bounds in option prices. The impact of these errors is pronounced and biased estimates of the trade-based variance are the result. So Harvey and Whaley (1991) document that neglecting the bid/ask spread and nonsimultaneous prices produces spurious autocorrelation in IVs.

The extent of these measurement errors is even stronger in a small market. But the focus on a small exchange raises a number of additional issues. In Vienna at-the-money options are less frequently traded than in the major US markets and the number of transactions also varies strongly over time, as there are days with very low turnover and periods where trading is quite active. All these observations call for the use of a more robust method for obtaining the IVs than the simple inverting of short-maturity ATM options. The procedure introduced by Lamoureux and Lastrapes (1993) is particularly well-adapted to our situation: Instead of using only a single closing option price the daily implied volatility is estimated from all available intra-daily transaction prices in a two-step method. First the pricing error e_i for option i is computed as follows:

$$e_i = p_i(Div_i, r_t^F, T_i, h_t, S_i) - p_i^o \tag{1}$$

Here p_i is the Black Scholes price, with Div_i as the dividend, r_t^F as the risk-free rate with maturity matched to the expiry date of the option, T_i is the maturity of option i, h_t is the volatility, S_i is the price of the underlying stock, and p_i^o is the observed transactions price from the options market. In the second step the loss function defined as the sum of the squares of e_i is minimized for all options $i \in \mathcal{I}$ on each day. This function depends only on h_t:

$$Loss(h_t) = \sum_{i \in \mathcal{I}} e_i^2(h_t) \tag{2}$$

With this two - step procedure we obtain the implied volatility for day t. By including more observations than just the closing quote this approach reduces the impact of measurement errors.

3.2 Time-Series based Volatility Models

In addition to the implied volatility we use four returns - based predictors of the daily out of sample volatility h_t: The "naive" moving average model, the Exponentially Weighted Moving Average (EWMA), Stochastic Volatility and GARCH. These models are all estimated on the daily returns computed as the first differences of log closing prices. In all time series models the mean of returns is set to zero[5].

The parameters are estimated using a rolling window of three years (738 daily observations). This window is moved forward every 20 days by adding 20 new observations and deleting the oldest 20 observations. The models are then used to predict out of sample volatility estimates for a time period corresponding to the maturity of the shortest option contracts.

3.2.1 The Equally Weighted Moving Average model

The benchmark for measuring the performance of any volatility model is the updated sample variance, i.e. the moving average of length n:

$$h_t = \frac{1}{n} \sum_{i=t-n}^{t-1} r_i^2 \tag{3}$$

with: $r_t \ldots$ compound return on day t, i.e. $r_t = ln(P_t/P_{t-1})$

We examine a MA specification with n = 30 days, termed MA(30). The predicted volatility during the maturity of the option is set equal to the variability during the last 30 days. So it is assumed that the future volatility equals the current estimate of past variability. The sample period roughly equals the average maturity of the options.

[5]The p-values of the t-statistics for the hypothesis mean = 0 are all larger than 0.02.

3.2.2 The Exponentially Weighted Moving Average model

The main drawback of the above version of the moving average model is that it allocates equal weight to all daily returns in the sample. So the MA(30) neglects the stronger impact of recent innovations. Therefore it can not reproduce the stylized fact of volatility clustering. This drawback is lifted by the exponentially weighted moving average:

$$h_t = (1-\lambda)\sum_{i=0}^{\infty}\lambda^i r_i^2 = \lambda h_{t-1} + (1-\lambda) r_{t-1}^2 \tag{4}$$

We estimate the decay factor λ by optimizing the log-Likelihood functions of the five stocks. From our estimation results we can also document to what extent the choice of RiskMetrics^{TM} (i.e. $\lambda = 0.94$) is appropriate in Vienna. Given the weight the EWMA forecasts for k days ahead are computed as follows:

$$\begin{array}{ll} k=1: & h_{t+1} = \lambda h_t + (1-\lambda) r_t^2 \\ k>1: & h_{t+k} = h_{t+k-1} \end{array} \tag{5}$$

3.2.3 The GARCH model

GARCH is a commonly used volatility model. It was introduced by Engle (1982) and generalized by Bollerslev (1986):

$$\begin{array}{rl} r_t \sim & N(0, h_t) \\ h_t = & a_0 + a_1 r_{t-1}^2 + a_2 h_{t-1} \end{array} \tag{6}$$

The GARCH parameters are also estimated by maximum likelihood from the rolling window. In order to avoid negative variances all coefficients in equation 6 must be nonnegative. The equations for the k-day predictions are given below:

$$\begin{array}{ll} k=1: & h_{t+1} = a_0 + a_1 r_t^2 + a_2 h_t \\ k>1: & h_{t+k} = a_0 + (a_1 + a_2) h_{t+k-1} \end{array} \tag{7}$$

In the long run GARCH forecasts converge to the unconditional variance. Comparing GARCH with EWMA we note many similarities. In both models today's volatility is deterministic given yesterday's return and variance. In contrast to EWMA, GARCH includes a constant term in its variance specification and it does not assume a persistent variance from the outset. EWMA however has a nonstationary variance process by construction, i.e. there is a unit root in second moments. This difference is clearly visible when comparing equations 5 and 7. We observe that the EWMA prediction for future volatility is a random walk.

3.2.4 The Stochastic Volatility model

Unlike the deterministic GARCH and EWMA volatilities the SV model contains innovations in both first and second moments. In its continuous-time specification it is used in option pricing, notably by Hull and White (1987). For the estimation on daily returns we use the following discrete-time specification introduced by Taylor (1986):

$$\begin{array}{ll} r_t = \sigma \exp(\xi_t/2)\varepsilon_t & \varepsilon_t \sim N(0,1) \\ \xi_t = \phi\xi_{t-1} + \eta_t & \eta_t \sim N(0,\sigma_\eta^2) \end{array} \tag{8}$$

This parameterization consists of the product process for returns and an AR(1) for the logarithm of volatility. In contrast to GARCH the SV specification has two innovations, ε_t and η_t , which are uncorrelated. Due to these disturbances in both first and second moments the model can not be estimated by means of maximum likelihood. Several methods for estimation have been proposed: Quasi Maximum Likelihood (QML) and Kalman filtering used by Harvey and Shephard (1996), Markov Chain Monte Carlo (Kim et al. (1996)) or the Efficient Method of Moments studied by Gallant et al. (1994). We have chosen the QML approach which is relatively simple to implement. QML is a two-step method. In the first step we conduct the following transformation on the model above:

$$log(r_t^2) = \xi_t + log(\varepsilon_t^2) + log(\sigma^2) \tag{9}$$

Then the Kalman filter[6] is used to extract the unobservable series ξ_t. The

[6] See Hamilton (1994), chapter 13 for details.

QML estimation delivers the parameters σ, σ_η^2 and ϕ. This approach is again implemented on the rolling window of daily returns as it is the case with GARCH and the moving averages. Because of the two step approach tests based on the likelihood function are not valid. So a direct comparison of the results from GARCH estimated by ML and SV estimated by Kalman/QML by means of Likelihood Ratio tests is rendered impossible.

The volatility predictions are computed from:

$$\begin{array}{ll} k=1: & h_{t+1}=h_t+exp(\phi) \\ k>1: & h_{t+k}=h_{t+k-1}+exp(\phi^k) \end{array} \quad (10)$$

3.3 Performance Measures

We use four criteria to evaluate the out-of-sample fit: Root Mean Squared Error (RMSE),Root Mean Squared Percentage Error (RMSPE), Root Mean Squared Logarithmic Error (RMSLE) and the R^2 from a regression on realized volatility. From these four measures we obtain the rankings of the performance of trade- and time series - based models. We define the realized volatility s_t as the mean of absolute daily returns over the forecasting horizon, i.e. the remaining lifetime of the short - maturity options.

$$\begin{array}{ll} RMSE = & \sqrt{\frac{1}{n}\sum_{t=1}^{n}(\sqrt{h_t}-s_t)^2} \\ RMSLE & = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(ln\left(\sqrt{h_t}\right)-ln\left(s_t\right)\right)^2} \\ RMSPE & = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\frac{(\sqrt{h_t}-s_t)^2}{s_t^2}} \end{array} \quad (11)$$

with:

s_t	realized volatility
$\sqrt{h_t}$	volatility forecast

In econometrics the most common error measure is the mean squared error, because it equals the sum of the squared bias and the variance of the estimator. To interpret the relative size of the forecast errors the Root Mean Squared Percentage Error (RMSPE) is more appropriate. When it is multiplied by 100 the RMSPE shows the percentage errors from the alternative models. These percentage values can be related to transactions costs to judge on the economic significance of the forecast errors. On this basis the pricing effects of using the different models become more visible. The third variant of mean squared error measures is the Root Mean Squared Logarithmic Error (RMSLE). Its use is motivated by the following observation: Given that volatility is always nonnegative, a skewed distribution of forecast errors is produced. This statistical property of the errors is neglected by RMSE and RMSPE, but it is taken into consideration by the RMSLE.

A joint drawback of all three criteria is the absence of significance levels. So we have no clear rule for when the error measures are insignificantly close. This is the reason for applying an alternative criterion, analyzed in Andersen and Bollerslev (1997). Instead of searching for the most accurate forecast, i.e. the method with the lowest RMSE, the focus here is on the information content and bias of forecasts. So the realized volatility is regressed on a constant and predicted volatility,

$$s_t = c_0 + c_1\sqrt{h_t} + u_t \tag{12}$$

where $\sqrt{h_t}$ denotes the volatility forecast. In the absence of bias, $E(s_t) = \sqrt{h_t}$, c_0 should be equal to zero and c_1 should be equal to unity. We use t-tests to evaluate the potential bias and the R^2s to measure which forecast explains most of the realized volatility. Equation 12 is estimated by ordinary least squares with the covariance matrix of White (1980), making the standard errors robust against heteroscedasticity and autocorrelation. This correction is necessary due to the overlapping observations. The advantage of the OLS method is that we can gain information on the trade-off between the variance and the bias of a forecast and the rankings can be compared in a simple manner. The final step is the test for orthogonality, where the information content of the competing models is evaluated. If e.g. GARCH is the optimal forecast its information set should be orthogonal to all alternatives, as the remaining models do not add any information to the dominant method. A test procedure is readily available from extending the two variable OLS regression defined above:

$$s_t = c_0 + c_1\sqrt{h_t} + c_2\sqrt{v_t} + u_t \tag{13}$$

Here v_t is an alternative model and the rest of the parameters are defined as in equation 12. This encompassing regression procedure enables us to test directly whether GARCH, SV or EWMA contain additional information, which is not in the IVs or vice versa. The t-statistics and the R^2 indicate whether the alternative historical models provide additional information not contained in the trade-based forecast of the IV. Alternatively this procedure can be interpreted as the construction of the optimal forecast. Then the coefficients from the regressions serve as the weights of the volatilities in a combined predictor.

4 Empirical Results

To visualize the estimation results Figure 1 shows the annualised out-of-sample volatility forecasts from EWMA, GARCH, the implied volatility, MA(30) and SV again for CA. First we note that volatility lies on average between 15% and 30% per annum. There exist wide divergences in the behaviour of the time series. There is evidence of a rise in the conditional variances, as many series indicate a positive trend. The volatility implied by the option prices mostly lies below the volatility series from the time series models and it has larger peaks than the returns - based models with a maximum close to 60%. This value is much larger than for the returns based models, where the maxima lie between 30 and 40%. The persistence of these shocks is low and the IV quickly reverts to its earlier level. These peaks do not always coincide with those of the returns-based models. So we note that the impact of shocks is different for the IV and for the historical measures. The two moving averages, MA(30) and EWMA, share many similarities. Thus the different decay factors seem to produce relatively small differences. Finally, it is observed that the stochastic volatility moves rather in a continuous manner and has a behaviour similar to GARCH.

Now we turn to the measurement of the out-of-sample fit. Table 5 contains Root Mean Squared Error, Root Mean Squared Percentage Error and Root Mean Squared Logarithmic Error. Of primary interest are the RMSPEs because they quantify the size of the forecast errors under the different models in percentage values. The tables indicate that the average forecast errors are between 30% and 75%. For the purpose of comparison we note that Figlewski (1997) obtains values around 50% for US asset returns. For our sample we can improve these results clearly for CA, but not for the five stocks. Altogether, the RMSPE ranks four times the MA(30) and once GARCH as best predictor. CA has relatively low error measures (29% to 36%), whereas Wienerberger achieves comparatively high values (57% and 75%). These observations indicate strong performance differences across the five stocks. So the out-of-sample fit does not show a uniform behaviour across the stocks in our sample. Across all stocks, MA(30) is most frequently chosen (nine times), but EWMA, SV and GARCH are also selected. Thus we observe that the more complex time-series based estimators such as GARCH and SV do not achieve a clear dominance relative to the simpler methods such as the MA(30) and EWMA.

To analyse the individual stocks in detail we start with CA. Here the rankings differ not only across the five stocks, but also across the three criteria. So each error measure chooses a different model as optimal: For the RMSE and RMSLE it is the short-horizon moving average, but for the RMSPE GARCH dominates. At the bottom of the rankings, there are also mixed results: For both RMSE and RMSLE the SV forecast has the worst fit. In contrast the RMSPE indicates the inferiority of IV. As another example we look at EVN. Here RMSE and

RMSLE select the exponentially weighted moving average (EWMA), but the RMSPE chooses the MA(30). Across all stocks, the implied volatility is never dominant. Thus according to Root Mean Squared Error, Root Mean Squared Percentage Error and Root Mean Squared Logarithmic Error the predictors based on historical returns outperform those based on option prices.

Given the divergence in the findings across RMSPE, RMSLE and RMSE, it is difficult to decide which criterion is most informative. All three measures have advantages and disadvantages. From a theoretical perspective RMSE is the commonly used selection criterion in econometrics. However it assumes a symmetric distribution of forecast errors. By construction, the distribution of volatility forecast errors is skewed, because volatility is always nonnegative. This criticism is also valid for RMSPE, but not for RMSLE. On the other hand the RMSPE is less sensible to nonstationarity as it uses relative errors. So the advantage of the RMSPE relative to the RMSE measure is twofold. First it gives a percentage value for the error relative to the outcome and is thus more easily interpreted. Second it is more appropriate for nonstationary variables because the ratio removes the joint trends, therefore reducing the impact of I(1) drifting. As the EWMA contains a unit root by construction, this is an advantage for the RMSPE. However a problem for all three criteria is that the values are in many cases quite close and we lack standard errors to gauge the significance of these differences. These criteria also do not distinguish between the variance and the bias of the predictors.

For this distinction we turn to the regression-based evaluation. As mentioned earlier the realized volatility is regressed on a constant and the forecast. The results then give information both on the bias and also on the information content of the forecast. Table 6 and following show the coefficients together with their White (1980) standard errors, the bias test and the measure of determination, R^2. The slopes and constants indicate that almost all predictors are biased. There is one exception, namely the MA(30) for VER. However, this slope shows the wrong size and the non-rejection of the bias is only caused by the size of the OLS standard errors. This finding implies that all models contain systematic forecast errors. When we study the ranking according to the size of the bias, as measured by the F-test, we find that the IV has the lowest bias for two stocks and GARCH, EWMA, MA(30) in one case. To analyse an individual stock we again turn to CA. Here the significant biases in all forecasts are clearly evident. The slope of SV with a value of 2.13 is significantly larger than unity, whereas the reverse is the case for EWMA, MA(30) and IV. Also for EVN, OMV and WIE, SV has a slope larger than unity. The results for VER deviate from those for the other stocks. Here all slopes are less than unity. Also, the out-of-sample performance is in general lower than with the other four stocks. Overall, our primary result is the pronounced bias in both the trade- and returns-based predictors.

Despite its systematic error a forecast can still be useful for a risk manager if it contains valuable information about the realizations of the variance until the maturity of the option. Here the R^2s are the appropriate statistics. The values from the two-variable regressions document that our forecasts explain between zero percent and 46 percent of the variability of the realized volatility. Again an ambiguous picture emerges: For CA and VER the highest values are observed for the SV model. For EVN and OMV the exponentially weighted moving average dominates and for Wienerberger GARCH is ranked first. The R^2s of the IV are between zero and 25 percent. From the univariate perspective, these values indicate that the trade-based volatility is inferior to time-series based predictors in all five cases. Therefore, when only one model is chosen, then this specification should be either SV, GARCH or a moving average. So next we turn to the selection of combinations of predictors of volatility.

The final test of forecasting performance is whether the volatility expected by participants in the options market adds information to predictors using historical returns. So far, the univariate error measures indicate a poor performance of the trade-based volatility. The multivariate procedure allows for a combined volatility forecast to dominate, i.e. a predictor weighing IV and a returns-based volatility. In this test the Null hypothesis is orthogonality of information sets, i.e. all relevant information is contained in the time-series based methods. The results of the regressions including both a time series-based and the implied volatility are given in the lower half of each panel. The coefficients of the IV in the regressions indicate that the volatility contained in option prices adds information, because many coefficients for the IV in the regression differ significantly from zero. This result is also supported by the adjusted R^2s. Here the result of the earlier procedure is unchanged only for CA and WIE, where SV was selected. However for EVN, OMV and VER, the optimal predictor is a combination of the implied volatility and EWMA, SV and GARCH. Thus in contrast to the three error criteria the regression-based evaluation delivers clearer results: We can conclude that the implied volatility contains valid information, which is absent in returns-based predictors. So because of this rejection of the orthogonality in three cases, the best forecast for future volatility combines two models. The weights for this optimal forecasts can be taken from the regressions: e.g. for EVN the corresponding best forecast is 0.75 EWMA + 0.37 IV.

5 Conclusion

This paper has compared the variance implicit in the prices of options on stocks of the Vienna Stock exchange to forecasts using time series of past price changes. Overall five models were evaluated: Implied Volatility, EWMA, MA(30), GARCH and SV. Our estimate of the trade-based volatility is obtained from transactions data on options with a method which is relatively robust to illiquidity and other features of small markets. To compare the out-of-sample performance we use a

Table 5
Error Measures

	MA(30)	EWMA	GARCH	SV	IV
			PANEL A:CA		
RMSE	0.08012	0.09271	0.09378	0.11174	0.09446
RMSLE	0.29760	0.34027	0.34167	0.41699	0.34100
RMSPE	0.31361	0.30497	0.29321	0.36042	0.36902
			PANEL B: EVN		
RMSE	0.09021	0.07513	0.08080	0.10331	0.08872
RMSLE	0.39499	0.36186	0.39637	0.46306	0.41474
RMSPE	0.43191	0.51362	0.44303	0.51112	0.61426
			PANEL C: OMV		
RMSE	0.07134	0.06976	0.07216	0.08137	0.07401
RMSLE	0.34408	0.36726	0.38581	0.40163	0.38319
RMSPE	0.38471	0.57418	0.53746	0.43398	0.51867
			PANEL D: VER		
RMSE	0.06251	0.09826	0.07698	0.05762	0.07028
RMSLE	0.33490	0.44192	0.43312	0.32812	0.35992
RMSPE	0.33608	0.74149	0.46964	0.42692	0.49344
			PANEL E: WIE		
RMSE	0.08365	0.09174	0.09409	0.10978	0.09329
RMSLE	0.36758	0.40227	0.41419	0.49458	0.43707
RMSPE	0.57099	0.66652	0.64775	0.64368	0.75345

where

RMSE is the Root Mean Squared Error, $RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(\sqrt{h_t} - s_t)^2}$

RMSLE is the Root Mean Squared Log Error, $RMSLE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(ln\left(\sqrt{h_t}\right) - ln\left(s_t\right)\right)^2}$

RMSPE is the Root Mean Squared Percentage Error, $RMSPE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\frac{(\sqrt{h_t} - s_t)^2}{s_t^2}}$

with:

s_t	realized volatility, i.e. mean of absolute daily returns over the forecasting horizon.
$\sqrt{h_t}$	volatility forecast

number of statistical criteria. Our main result concerns the ranking of the models. We observe that the choice strongly depends on how the forecasting ability is measured. So among the five models no clear winner emerges. Among the stocks, there are pronounced performance differences. When we compare the relative out-of-sample fit of the four time series models, we find that the complex approaches (GARCH and SV) do not dominate the simpler moving average models. Our second result is the existence of systematic prediction errors. Hence, all five volatility predictors are significantly biased. When we analyse the importance of

Table 6
Regressions on realized volatility - CA

Constant	Slope on MA(30)	EWMA	GARCH	SV	IV	Bias Test	R^2
0.1083	0.6663					39.74	0.2158
(0.0531)*	(0.1753)+					[0.000]	
0.0866		0.7693				42.37	0.2146
(0.0457)		(0.1728)+				[0.000]	
-0.1428			1.7875			32.74	0.1834
(0.0937)			(0.3785)+			[0.000]	
-0.2503				2.1301		133.10	0.4653
(0.0419)*				(0.2039)+		[0.000]	
0.1094					0.6237	16.93	0.1046
(0.0663)					(0.2392)+	[0.000]	
0.0648	0.5742				0.2708		0.2314
(0.0423)	(0.1855)+				(0.1890)+		
0.0608		0.6779			0.1946		0.2218
(0.0430)		(0.1956)+			(0.1882)+		
-0.1349			1.5019		0.2368		0.1939
(0.0895)			(0.2001)+		(0.1245)+		
-0.2510				2.1214	0.0114		0.4654
(0.0370)*				(0.3410)+	(0.2353)		

The table shows the results from regressing the alternative daily volatility forecasts $\sqrt{h_t}$ on the realized standard deviation s_t; constant and slope have the White (1980) standard errors in brackets, in the univariate regressions * signifies constant significantly different from 0 at 5%, + signifies slope significantly different from 0 at 5%; for the Bias Test [H_0: Constant = 0 and Slope = 1] the brackets gives the corresponding marginal significance levels.

the trade-based volatility we find that it contains exclusive information, which the predictors based on historical returns do not contain. Hence for our sample the optimal volatility forecast combines the predictions from historical returns and from option prices.

Therefore we can conclude that special care has to be taken when selecting a volatility model for either risk management or derivatives pricing in smaller markets. Current time series based models as well as implied volatilities are not capable of explaining future variability of stock returns. For future research we find two areas of interest. First there is demand for extending current volatility specifications. Here in particular the optimal combination of forecasts from

Table 7
Regressions on realized volatility - EVN

Constant	Slope on					Bias Test	R^2
	MA(30)	EWMA	GARCH	SV	IV		
0.0537	0.8013					60.66	0.3745
(0.0157)*	(0.1147)+					[0.000]	
0.0223		0.8683				56.62	0.4232
(0.0157)		(0.1316)+				[0.000]	
-0.0593			1.3173			50.64	0.3762
(0.0314)			(0.2046)+			[0.000]	
-0.1829				2.2339		35.07	0.3501
(0.0676)*				(0.4103)+		[0.000]	
0.0301					0.8277	9.22	0.1979
(0.0489)					(0.3276)+	[0.000]	
0.0001	0.6774				0.3711		0.4054
(0.0284)	(0.1091)+				(0.1805)+		
-0.0318		0.7519			0.3794		0.4573
(0.0261)		(0.1340)+			(0.1604)+		
-0.0969			1.1139		0.3782		0.4086
(0.0255)*			(0.2270)+		(0.1544)+		
-0.2108				1.8571	0.4503		0.3988
(0.0611)*				(0.2647)+	(0.2026)+		

Table 8
Regressions on realized volatility - OMV

Constant	Slope on					Bias Test	R^2
	MA(30)	EWMA	GARCH	SV	IV		
0.0825	0.6073					118.89	0.2060
(0.0184)*	(0.0778)+					[0.000]	
0.0582		0.6592				2.54	0.2089
(0.0069)*		(0.0000)+				[0.010]	
-0.0282			1.0688			57.94	0.1505
(0.0275)			(0.1602)+			[0.000]	
-0.1401				1.8643		10.17	0.1907
(0.0954)				(0.6356)+		[0.000]	
0.0622					0.6165	12.91	0.0882
(0.0465)					(0.2309)+	[0.000]	
0.0570	0.5493				0.1726		0.2111
(0.0385)	(0.0613)+				(0.1510)+		
0.0197		0.5825			0.2594		0.2218
(0.0349)		(0.0000)+			(0.1706)+		
-0.0848			0.9124		0.4266		0.1896
(0.0219)*			(0.1053)+		(0.2262)+		
-0.2459				1.7976	0.5664		0.2650
(0.0708)*				(0.5777)+	(0.2084)+		

Table 9
Regressions on realized volatility - VER

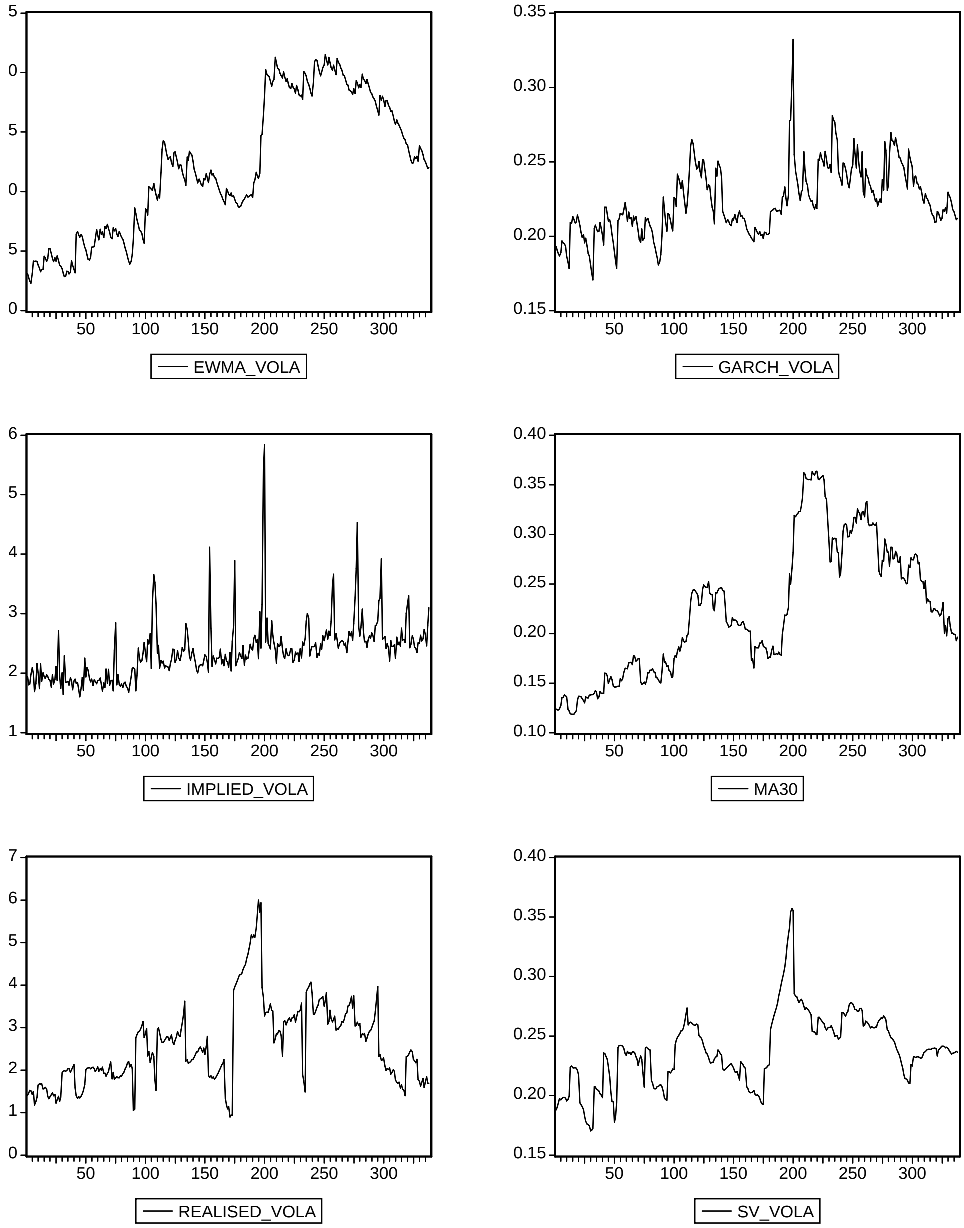
Constant	Slope on					Bias Test	R^2
	MA(30)	EWMA	GARCH	SV	IV		
0.2216	-0.2525					0.02	0.0274
(0.0585)*	(0.3105)					[0.900]	
0.1788		0.0134				22.36	0.0003
(0.0177)*		(0.0566)				[0.000]	
0.0147			0.7210			14.11	0.0739
(0.0650)			(0.2609)+			[0.000]	
0.0431				0.8935		3.48	0.0281
(0.0931)				(0.5946)		[0.060]	
0.1727					0.0454	14.79	0.0015
(0.0111)*					(0.0648)	[0.000]	
0.2120	-0.2544				0.0509		0.0294
(0.0542)*	(0.3137)+				(0.0739)+		
0.1703		0.0124			0.0447		0.0018
(0.0198)*		(0.0573)			(0.0647)+		
0.0083			0.7185		0.0362		0.0750
(0.0698)			(0.2624)+		(0.0564)		
0.0331				0.8971	0.0485		0.0299
(0.0923)				(0.5996)+	(0.0594)+		

Table 10
Regressions on realized volatility - WIE

Constant	Slope on					Bias Test	R^2
	MA(30)	EWMA	GARCH	SV	IV		
0.0791	0.6727					85.55	0.2627
(0.0374)*	(0.1182)+					[0.000]	
0.0675		0.7067				98.18	0.2741
(0.0365)		(0.1134)+				[0.000]	
0.0726			0.6905			39.81	0.2249
(0.0506)			(0.1695)+			[0.000]	
-0.0590				1.4162		170.39	0.5053
(0.0173)*				(0.1185)+		[0.000]	
0.0302					0.8963	66.95	0.2512
(0.0422)					(0.1464)+	[0.000]	
0.0158	0.4323				0.5368		0.3193
(0.0311)	(0.1598)+				(0.1398)+		
0.0105		0.4668			0.5136		0.3250
(0.0316)		(0.1532)+			(0.1394)+		
0.0115			0.3817		0.5999		0.2925
(0.0426)			(0.2079)+		(0.1487)+		
-0.0629				1.3697	0.0615		0.5060
(0.0250)*				(0.1127)+	(0.2057)		

returns and from option prices is of considerable interest. Second better model selection criteria might eliminate ambiguity in performance rankings.

Figure 1
Volatility series for CA



References

K. Adjaoute, M. Bruand, and R. Gibson, On the Predictability of Stock Market Volatility: Does History Matter?, *European Financial Management*, 4(3): 293–319, 1999.

T. Andersen and T. Bollerslev, Answering the Skeptics: Yes standard volatility models do provide accurate forecasts, mimeo, 1997.

D. Bates. Testing Option Pricing Models. In C.S. Maddala and C.R. Rao, editors, *Handbook of Statistics*, volume 14, pages 567–612. 1996.

F. Black and M. Scholes, The pricing of options and corporate liabilities, *Journal of Political Economy*, 8: 1637–1659, 1973.

T. Bollerslev, Generalized Autoregressive Conditional Heteroscedasticity, *Journal of Econometrics*, 31: 307–327, 1986.

L. Canina and S. Figlewski, The Informational Content of implied volatility, *Review of Financial Studies*, 6: 659–682, 1993.

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

J. Duan, The GARCH Option Pricing Model, *Mathematical Finance*, 5: 13–32, 1995.

R. Engle, Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of U.K. Inflation, *Econometrica*, 50: 987–1008, 1982.

S. Figlewski, Forecasting Volatility, *Financial Markets, Institutions and Instruments*, pages 1–88., 1997.

J. Fleming, The quality of market volatility forecasts implied by S&P 100 index options, *Journal of Empirical Finance*, 5: 317–345, 1998.

A. Gallant, D. Hsieh, and G. Tauchen, Estimation of Stochastic Volatility models with diagnostics, mimeo, 1994.

J. Hamilton, Time Series Analysis. Princeton University Press., 1994.

A.C. Harvey and N. Shephard, Estimation of an asymmetric model of asset prices, *Journal of Business and Economic Statistics*, 14: 429–434, 1996.

C.R. Harvey and R.E. Whaley, S&P 100 Index Option Volatility, *Journal of Finance*, 46: 1551–1561, 1991.

S. Heston, A closed form solution for options with stochastic volatility with applications to bond and currency options, *Review of Financial Studies*, 63: 27–44, 1993.

R.C. Heynen and H.M. Kat, Volatility Predictions: A Comparison of the SV, GARCH(1,1), and EGARCH Models, *Journal of Derivatives*, Winter: 50–65, 1994.

J. Hull and A. White, The pricing of options on assets with stochastic volatility, *Journal of Finance*, 42: 281–300, 1987.

P. Jorion, Predicting Volatility in the Foreign Exchange Market, *Journal of Finance*, 50: 507–528., 1995.

D. Kim, N. Shephard, and S. Chib, Stochastic Volatility Likelihood Inference and Comparison with ARCH models, mimeo, 1996.

C. Lamoureux and W. Lastrapes, Forecasting Stock Return Variances: Towards an Understanding of stochastic Implicit Volatilities, *Review of Financial Studies*, 6: 293–326, 1993.

R.J. Mahieu and P. Schotman, An Empirical Application of SV models, *Journal of Applied Econometrics*, 13: 333–360, 1998.

M. Pagano, A.A. Röell, and J. Zechner, The Geography of Equity Listing: Why Do European Companies List Abroad?, mimeo, 2000.

F. Palm. GARCH Models of Volatility. In C.S. Maddala and C.R. Rao, editors, *Handbook of Statistics*, volume 14, pages 209–239. 1996.

M. Scheicher, Nonlinear Dynamics: Evidence for a Small Stock Exchange, *Empirical Economics*, 24: 45–60, 1999.

S.J. Taylor, Modeling Financial Time Series. Wiley, 1986.

H. A. White, Heteroscedastic-Consistent Covariance Matrix and a Direct Test for Heteroscedasticity, *Econometrica*, 48: 421–448, 1980.

J. Wiggins, Option values under stochastic volatility: Theory and empirical estimates, *Journal of Financial Economics*, 19: 351–377, 1987.