Financial Analysts Journal
Volume 61 • Number 1
©2005, CFA Institute

FAJ

Practical Issues in Forecasting Volatility

Ser-Huang Poon and Clive Granger

A comparison is presented of 93 studies that conducted tests of volatility-forecasting methods on a wide range of financial asset returns. The survey found that option-implied volatility provides more accurate forecasts than time-series models. Among the time-series models, no model is a clear winner, although a possible ranking is as follows: historical volatility, generalized autoregressive conditional heteroscedasticity, and stochastic volatility. The survey produced some practical suggestions for volatility forecasting.

Volatility forecasting plays an important role in investment, option pricing, and risk management. We conducted an extensive review of the volatility-forecasting research in the last 20 years (Poon and Granger 2003) and provide here a summary and update of our findings. The definition of volatility we used is the standard deviation of returns. The assets studied in the 93 articles surveyed included stock indexes, stocks, exchange rates, and interest rates from both developed and emerging financial markets. The forecast horizon ranged from one hour to one year (a few exceptions extended the forecast horizon to 30 months and to five years).

We review three main categories of time-series model—namely, historical volatility, models in the autoregressive conditional heteroscedasticity (ARCH) class, and stochastic volatility models—as well as forecasting based on implied volatility derived from option prices. We present here a description of these models, a summary of our survey results, and a discussion of the characteristics of market volatility that affect the choice of model, common objectives of volatility forecasting, and the impact of outliers. Finally, we provide some practical advice on volatility forecasting.

Types of Volatility Models

The four types of volatility-forecasting methods we surveyed are historical volatility (HISVOL), ARCH models, stochastic volatility, and option-implied volatility.

The HISVOL model is

$$\hat{\sigma}_t = \phi_1\sigma_{t-1} + \phi_2\sigma_{t-2} + \ldots + \phi_\tau\sigma_{t-\tau}, \qquad (1)$$

Ser-Huang Poon is professor of finance at the University of Manchester, United Kingdom. Clive Granger is professor emeritus and research professor at the University of California at San Diego.

where

$\hat{\sigma}_t$ = expected standard deviation at time t

ϕ = the weight parameter

σ = historical standard deviation for periods indicated by the subscripts

This group includes random walk, historical averages, autoregressive (fractionally integrated) moving average, and various forms of exponential smoothing that depend on the values of ϕ, the weight parameter.

The second group is the ARCH model and its various extensions, including the nonlinear ones:

$$r_t = \mu + \varepsilon_t, \qquad (2a)$$

where

r_t = return of the asset at time t

μ = average return

ε_t = residual returns, defined as

$$\varepsilon_t = \sqrt{h_t}z_t, \qquad (2b)$$

where z_t is standardized residual returns and h_t is conditional variance, defined as

$$h_t = \omega + \sum_{j=1}^{p}\beta_j h_{t-j} + \sum_{k=1}^{q}\alpha_k\varepsilon_{t-k}^2, \qquad (2c)$$

in which

ω = a constant term

p = number of autoregressive terms

j = order of the autoregressive term

β = autoregressive parameter

q = number of moving-average terms

k = order of the moving-average term

α = moving-average parameter

The stochastic volatility (SV) model is defined as

$$r_t = \mu + \varepsilon_t, \qquad (3a)$$

with

$$\varepsilon_t = z_t \exp(0.5h_t) \qquad (3b)$$

and

$$h_t = \omega + \sum_{j=1}^{p} \beta_j h_{t-j} + \upsilon_t, \tag{3c}$$

where υ_t is an innovation term. The variables υ_t and z_t could be correlated.

The fourth type of model deals with option-implied standard deviation (ISD) based on the Black–Scholes (1973) model and various generalizations. If g denotes the option-pricing model and c is the price of the option, then

$$c = g\,(S, X, \sigma, R, T), \tag{4}$$

where

S = price of the underlying asset
X = exercise price
σ = volatility
R = risk-free interest rate
T = time to option maturity

The ISD is the value that causes the right-hand side of Equation 4 to equal the market price of option c.

In Equation 1, the historical volatilities—that is, $\sigma_{t-1}, \sigma_{t-2}, \ldots, \sigma_{t-\tau}$—have to be calculated somehow from historical returns before the volatility model can be estimated. The various ways of calculating these historical volatilities and the different lengths of sample data used can lead to very different volatility forecasts. Recent research shows that daily realized volatility calculated from intraday squared returns measured at 5-minute or 15-minute intervals produces the best results.

The models given in Equations 2 and 3 are similar in being based on fitting the return distribution. This characteristic is convenient for the user because daily returns are available for many financial time series. The disadvantage of such an approach is that the volatility structure is then constrained by the choice of return distribution. For example, $\sum_{j=1}^{p} \beta_j + \sum_{k=1}^{q} \alpha_k$ should not exceed 1 in the ARCH model (Equation 2). The SV model (Equation 3) is more flexible than the ARCH model because of the second innovation term, υ_t. But the introduction of υ_t makes direct inference of ε_t much more complex. Limited research findings published to date provide no clear evidence to indicate that SV provides better forecasts than HISVOL or ARCH.

Option-implied volatility (Equation 4) works in a way that is completely different from the three time-series models. Technically, such information as historical returns and historical volatility is not needed. On the assumption that option-pricing function g is correct, a single option price is sufficient to produce an estimate of future volatility. Option market prices appear to have a premium, however, over Black-Scholes prices. Hence, Black–Scholes ISD tends to be higher than actual volatility. To overcome this bias, historical volatility is used for calibration, as follows:

$$\sigma_{t+1} = \alpha + \beta\, ISD_t + e_t \text{ with } t = 1, 2, \ldots, T-1 \tag{5a}$$

and

$$\hat{\sigma}_{T+1} = \alpha + \beta\, ISD_T, \tag{5b}$$

where α and β are regression parameters and $\hat{\sigma}_{T+1}$ is the volatility forecast at $T + 1$. The time t option price and ISD contain volatility information on the future up to option maturity.

Volatility-Forecasting Contests

In our review of the results in 93 volatility studies, we excluded all the papers that had no forecasting content and the papers with forecasts that are not out of sample. **Table 1** summarizes outcomes for the 66 papers that provided pairwise comparisons (GARCH stands for generalized autoregressive conditional heteroscedasticity). The first column should be read as follows: A > B means Model A performed better than Model B.

Table 1. Summary of Pairwise Outcomes for 66 Studies

Comparison Outcome	Number of Studies	Percentage of Studies
HISVOL > ARCH	22	56
ARCH > HISVOL	17	44
HISVOL > ISD	8	24
ISD > HISVOL	26	76
GARCH > ISD	1	6
ISD > ARCH	17	94
SV > HISVOL	3	
SV > ARCH	3	
ARCH > SV	1	
ISD > SV	1	

The overall ranking suggests that ISD provides the best forecasts, followed by HISVOL and GARCH with roughly equal performance, although HISVOL may perform somewhat better. The number of studies involving SV is so small that we could not make any clear statement about the SV model.

The success of the implied-volatility method should not be surprising because these forecasts are based on a larger and timelier information set. Options are written on limited classes of assets, however, and are traded in only a handful of exchanges. For example, equity stocks in many emerging markets are important components of an international equity portfolio but many of these stocks and stock market indexes have no listed option contracts. So, time-series models, although inferior to option-implied models, will continue to play an important role in volatility forecasting.

Among the 93 papers, 17 studies compared alternative versions of ARCH. Among the 17 studies, the more general GARCH clearly dominated ARCH. In general, models that incorporated volatility asymmetry, such as EGARCH ("E" for "exponential") and GJR-GARCH ("GJR" for Glosten, Jagannathan, and Runkle 1993), performed better than GARCH, but certain specialized specifications, such as fractionally integrated GARCH and regime-switching GARCH did better in some studies.

An important question that has not yet been addressed is: How well do the volatility-forecasting models complement each other cross-sectionally and through time? Different methods may be capturing the information set differently, and which method is superior may depend on market conditions. Unfortunately, little research has been done on the performance of combined volatility forecasts. Only 3 of the 93 papers surveyed evaluated a combination of forecasts. Two studies found it to be helpful, but another did not.

Also rarely discussed in the 93 papers is whether one method is significantly better than another. The forecast evaluation criteria in the papers often bear no relation to the objectives of volatility forecasting as we outline them later. Thus, although we can suggest that a particular method of forecasting volatility is the best, we cannot state that the benefits of a method outweigh the costs of using it rather than some simpler approach.

The comparisons we have made here are broadly based and brush aside many finer points. For example, the papers reviewed did not all study identical assets over the same sample period or adopt the same forecast horizon. Moreover, the survivorship bias in the publication process inevitably leads to some studies being conducted simply to support the viewpoint that a particular method is useful (that is, the paper might not have been submitted or accepted for publication if the required result had not been reached). This bias is one of the obvious weaknesses of a study such as ours.

Characteristics of Financial Market Volatility

Financial market volatility has a number of characteristics that are generally well cited in the literature. One of the facts is that volatility persists and clusters. This characteristic is illustrated in **Figure 1**. Panel A shows realized volatility of returns (calculated from cumulative intraday returns) on the S&P 500 Index for the period 1 February 1983 through 31 July 2003.[1] The S&P 500 volatility presented was truncated at 4 percent so that the series could be studied without the overwhelming dominance of three large values (10.0 percent on 19 October 1987, 14.3 percent on 20 October 1987, and 7.7 percent on 29 October 1997). Panel A shows that high-volatility days tend to group together and that the same is true for low-volatility days.

Panel B of Figure 1 presents the autocorrelation and partial autocorrelation coefficients of the first 1,000 lags of S&P 500 realized volatility.[2] Volatility persistence manifests itself in the autocorrelation coefficient, which remains significantly greater than zero after 1,000 lags. The partial autocorrelation coefficient approaches zero as lag length extends beyond 25. This strong persistence gives rise to the "long memory" effect, which we return to later.

Another important characteristic of the financial markets is volatility asymmetry, which is particularly prominent in the equity markets. **Figure 2** shows the impact of S&P 500 returns on S&P 500 volatility on the contemporaneous day and volatility on the following day. The scattergram is based on the following regression:

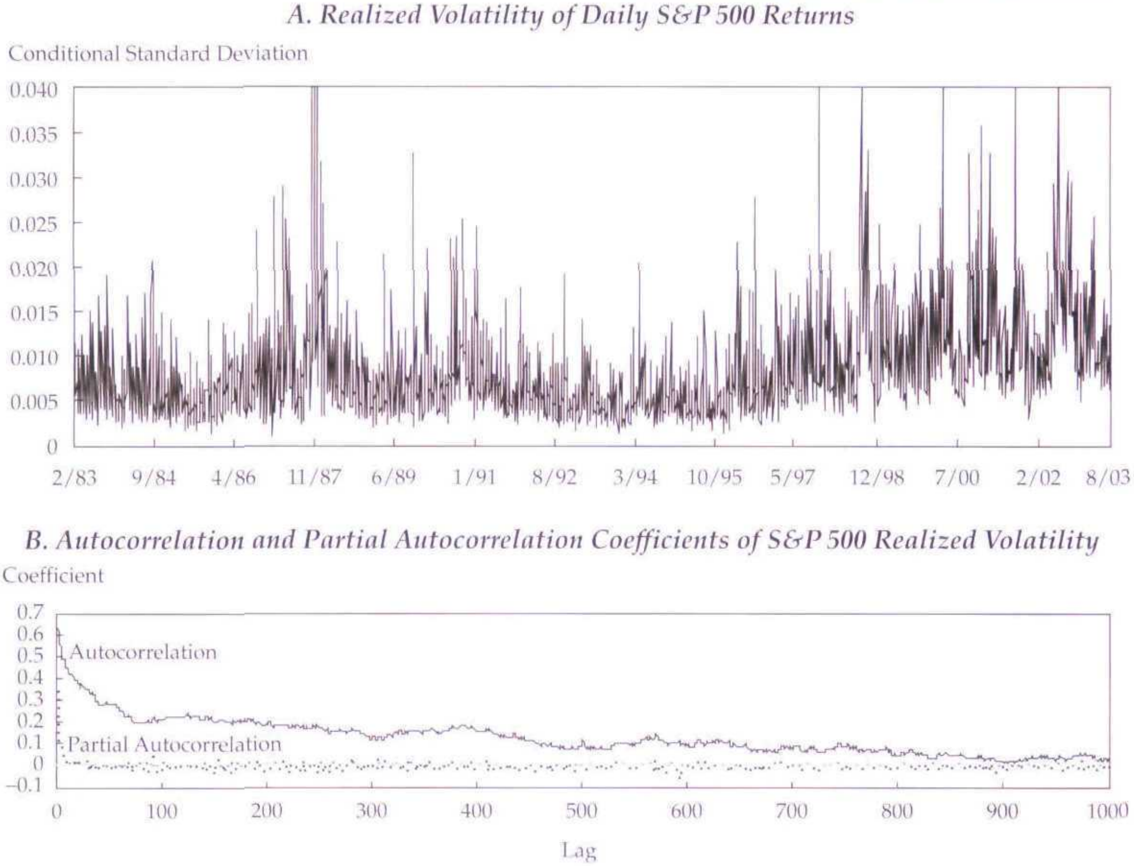
$$\begin{aligned}(1-\phi)\sigma_{RV} = {} & \omega + \beta_1 D_{t,1} r_t + \beta_2 (1 - D_{t,1}) r_t \\ & + \beta_3 D_{t,2} r_{t-1} + \beta_4 (1 - D_{t,2}) r_{t-1} \\ & + (1-\theta)\upsilon_t, \end{aligned} \tag{6}$$

where σ_{RV} is the realized volatility calculated from intraday S&P 500 returns, $D_{t,1}$ is a dummy variable that takes the value of 1 for $r_t < 0$ and 0 otherwise, and similarly, $D_{t,2}$ is 1 for $r_{t-1} < 0$ and 0 otherwise. With $\beta_1 > \beta_2$ and $\beta_3 > \beta_4$ in absolute terms, the impact of returns on volatility is clearly stronger in bear markets than in bull markets.

A similar phenomenon appears in interest rate series, but interest rates tend to be dominated by a level effect (whereby high volatility is associated with high interest rate levels and low volatility is associated with low interest rate levels).

Some stock markets have experienced shifts in volatility; an example is provided by the returns on the South Korean Stock Exchange Composite Index (KOSPI), shown in Panel A of **Figure 3**. The shift that is so visible coincided with the Asian crisis in 1997. A shift in volatility level can also be detected for some exchange rates and interest rates—possibly coinciding with the timing of policy changes. The impact on interest rates of the U.S. Federal Reserve's policy introduced in the 1980s can be clearly seen in the U.S. dollar one-month LIBOR, shown in Panel B of Figure 3.[3] But in the 300 or more financial time series that we have

Figure 1. Realized Volatility of S&P 500 Returns, 1 February 1983 through 31 July 2003



Note: With 5,168 observations in the S&P 500 realized volatility series, the standard error of (partial) autocorrelation coefficients is 0.0278.

encountered, we have found no steady linear upward trend in financial market volatility.

As noted, strong volatility persistence, or long memory, is another well-known fact about financial market volatility; it has been extensively discussed (see, e.g., *Journal of Econometrics* 1996, vol. 73, no. 1). Researchers have noticed that the autocorrelation of the function of returns, $|r_t|^d$ with $d > 0$, is slow to decay, particularly when $d = 1$ (Taylor 1986). **Table 2** presents the sum of autocorrelation coefficients of the first 1,000 lags for 20 selected financial time series and two simulated ARCH processes—GARCH(1,1) and GJR(1,1). Both simulated processes had specifications that produced strong volatility persistence. We used four types of daily volatility proxies: absolute return, $\Sigma\rho(|r|)$; squared return, $\Sigma\rho(r\wedge 2)$; logarithm of absolute return, $\Sigma\rho(\ln|r|)$; and trimmed absolute return, $\Sigma\rho(|Tr|)$. (Trimming is explained in the note to Table 2.) The logarithmic transformation and trimming procedure had the effect of reducing the impact of outliers, whereas taking the square of the returns amplified the influence of large values.

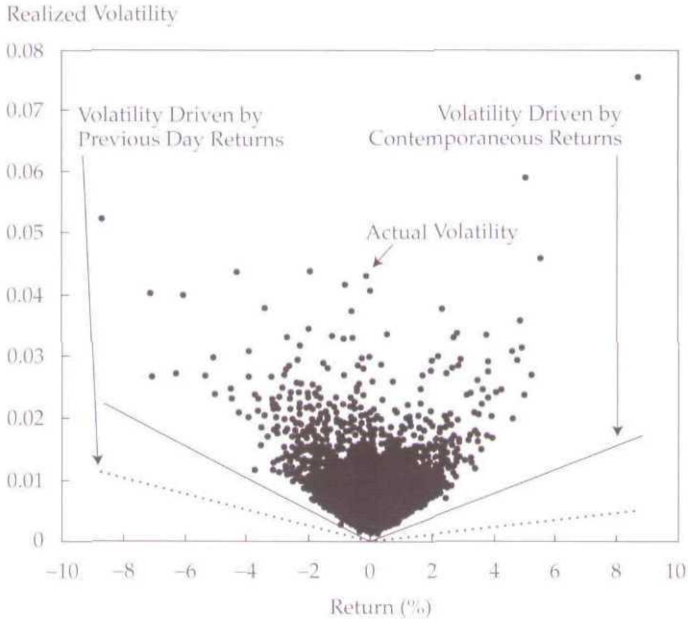
High autocorrelation values indicate long memory. Thus, Table 2 suggests that financial time series have far longer memories than do stationary GARCH and GJR processes. All the time-series volatility models were designed to capture volatility persistence. The stationary GARCH and GJR models had memories that were too short to fit the fact of long memory in volatility.

The fractionally integrated (FI) model is the only linear model that has a memory long enough to fit the empirical observations, and some researchers have found FI volatility models to forecast well.[4] The concern is that, even though FI models may match the characteristic of long memory, they may still not reflect the true volatility process.

The important question is: What is the economic explanation for such a long memory in financial market volatility? Do we expect financial markets and market participants to have memories as long as the memory implied in FI models?

At the time of this writing, researchers have found alternative nonlinear volatility models that will produce a long memory in absolute returns but

Figure 2. S&P 500 Volatility Asymmetry, 1 February 1983 through 31 July 2003



Parameter	Coefficient	Standard Error	t-Statistic
ω	0.0058	0.0003130	18.576
β_1	−0.0026	0.0000745	−34.663
β_2	0.0020	0.0000728	27.055
β_3	−0.0013	0.0000742	−17.504
β_4	0.0006	0.0000729	7.921
ϕ	0.9666	0.0045940	210.394
θ	−0.7603	0.0117180	−64.880

Note: Based on realized volatility, σ_{RV}. The adjusted R^2 of the regression is 62.74 percent, and the Durbin–Watson statistic is 1.97.

the volatility process has short-memory dynamics. These models include the break process in Granger and Hyung (2004), regime switching in Diebold and Inoue (2001), volatility components in Engle and Lee (1999), and the stochastic unit root process in Yoon (2003). These alternative models are intuitively appealing, and some of them provide a better fit to the empirical data than the FI models because of additional parameters. Whether they provide better forecasts is an empirical question.

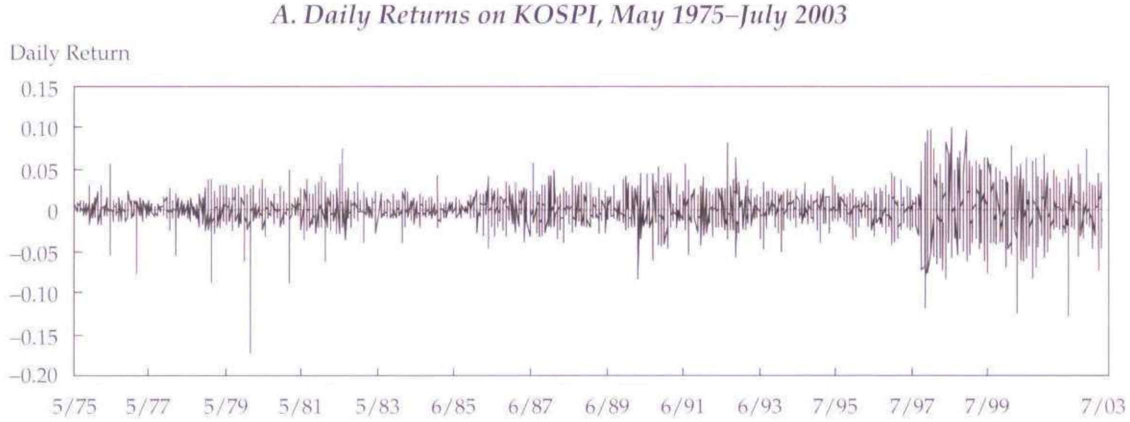
Objectives of Volatility Forecasting

The main reason for the prominent role that volatility plays in financial markets is that volatility is associated with risk and uncertainty, the key attributes in investing, option pricing, and risk management. Heteroscedasticity, a technical term for time-varying volatility, makes the estimation of asset-pricing relationships inefficient. Hence, econometric techniques are needed in controlling for heteroscedasticity in financial market modeling. ARCH and SV are useful in this pursuit because they are estimated on the basis of return distribution. ARCH models, in addition, are easy to implement.

Risk and Risk Management. Volatility is a measure for the second moment of a distribution. The first moment is the mean, the third is skewness, and the fourth, kurtosis. For a normally distributed variable, skewness is always 0 and kurtosis is always 3. So, the first two moments alone are sufficient statistics for summarizing the characteristics of the entire bell-shaped distribution. It is, therefore, convenient to equate return and risk to the first two moments of the return distribution, and indeed, this assumption is fundamental in Markowitz mean–variance portfolio theory and the capital asset pricing model.

Researchers have long noted, however, that financial asset returns are not normally distributed (Mandelbrot 1963; Fama 1965). Data collected since the 1960s show that stock market returns are usually negatively skewed and have high kurtosis. In the United States, for example, the excess kurtosis

Figure 3. Shifts in Volatility



(i.e., kurtosis in excess of 3) is 2.37 for 20-day returns and 35.58 for 1-day returns. If the period before 1985 is excluded, excess kurtosis is 44.07 and skewness is –2.1 for daily returns. Both figures are statistically different from zero. Similar patterns prevail in stock markets all over the world. They are clear evidence that stock market returns are anything but normal. ARCH standardized residuals are closer to normal but are still not normal. An asset-pricing model that takes into account higher moments and extreme events is needed.

If risk is defined as the possibility of negative returns and large losses, the lower quantiles are a more relevant risk measure than volatility because high volatility may be driven entirely by a large positive return. The industry practice of reporting value at risk (VAR) is, in fact, reporting the 1 percent quantile (or 0 if this figure is nonnegative). The 1 percent quantile for U.S. stock market returns is –2.57 percent, but the maximum one-day loss in the United States in the post-1985 data is 22.8 percent. Hong Kong's 1 percent quantile is –2.53 percent, which is smaller than the U.S. result, but the maximum one-day loss is a staggering 40.54 percent. Thus, the quantile is an incomplete description of the tail size. Expected shortfall is a better measure, and a good model of expected shortfall must involve extreme-value techniques.[5]

Option Pricing. An option represents a financial claim whose payoff is contingent on the occurrence of an uncertain event. For an equity call option, for example, the payoff will depend on how much the terminal stock price exceeds the exercise price. The risk-neutral valuation principle established by Black and Scholes means that the mean return on the stock is irrelevant and volatility is the most important factor in determining option prices. Hence, by observing option prices traded in the market, we can infer the market's view of future volatility over the option's maturity. Given the sophistication and efficiency of the financial markets in processing information, it is no surprise that option-implied volatility has been shown to possess stronger volatility-forecasting power than time-series models using only historical information. But there is a catch: Option-implied volatilities of different strike prices can be vastly different. The question that follows, then, is: Which of the implied volatilities should one use?

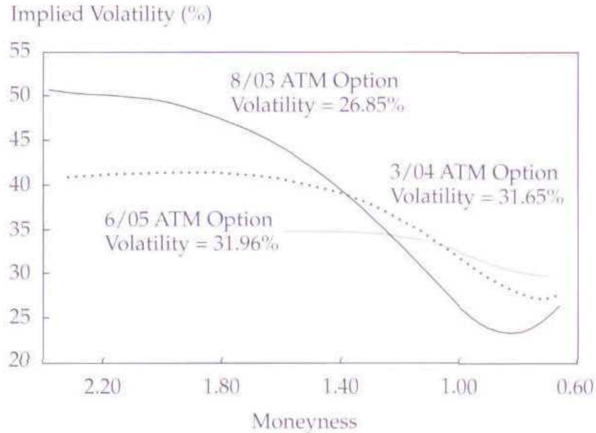
Table 2. Sum of Autocorrelation Coefficients of First 1,000 Lags in Selected Financial Time Series and Simulated ARCH Processes: Various Start Dates, Ending 22 July 2003

Data Series	N	$\Sigma\rho(\|r\|)$	$\Sigma\rho(r\hat{}2)$	$\Sigma\rho(\ln\|r\|)$	$\Sigma\rho(\|Tr\|)$
Stock market indexes					
S&P 500 Composite (U.S.)	9,676	35.687	3.912	27.466	40.838
DAX 30 Industrial (Germany)	9,634	75.571	37.102	41.890	79.186
NIKKEI 225 Stock Average (Japan)	8,443	89.559	23.405	84.257	95.789
CAC 40 (France)	8,276	43.310	17.467	22.432	46.539
FTSE All Share and FTSE 100	8,714	30.817	12.615	18.394	33.199
Average		54.989	18.900	38.888	59.110
Stocks					
Cadbury Schweppes	7,418	48.607	19.236	85.288	50.235
Marks and Spencer Group	7,709	40.635	17.541	67.480	42.575
Shell Transport	8,115	38.947	20.078	44.711	40.035
FTSE Small-Cap Index	4,437	25.381	3.712	35.152	28.533
Average		38.392	15.142	58.158	40.344
Exchange rates					
U.S. dollar/U.K. pound	7,942	56.308	24.652	84.717	57.432
Australian dollar/U.K. pound	7,859	32.657	0.052	72.572	48.241
Mexican peso/U.K. pound	5,394	9.545	1.501	13.760	14.932
Indonesian rupiah/U.K. pound	2,964	20.819	4.927	31.509	21.753
Average		29.832	7.783	50.640	35.589
Interest rates					
U.S. one-month Eurodollar deposits	8,491	281.799	20.782	327.770	331.877
U.K. interbank one-month rate	7,448	12.699	0.080	22.901	25.657
Venezuelan par Brady bond	3,279	19.236	9.944	32.985	19.800
Korean overnight call	2,601	54.693	12.200	57.276	56.648
Average		92.107	10.752	110.233	108.496
Commodities					
Gold, U.S. dollar/troy oz. (London Bullion Market), fixing, close	6,536	125.309	39.305	140.747	133.880
Silver, U.S. cents/troy oz. (London Bullion Market), fixing	7,780	45.504	8.275	88.706	52.154
Brent oil (one-month forward), U.S. dollar/barrel	2,389	11.532	5.469	9.882	11.81
Average		60.782	17.683	79.778	65.948
Average for all		54.931	14.113	65.495	61.555
Simulated GARCH					
Mean	10,000	1.045	1.206	0.478	1.033
Standard deviation		1.099	1.232	0.688	1.086
Simulated GJR					
Mean	10,000	1.945	2.308	0.870	1.899
Standard deviation		1.709	2.048	0.908	1.660

Note: *Tr* denotes trimmed returns, whereby all returns in the 0.1 percent tail are forced to take the 0.1 percent quantile value.

Figure 4 presents the implied volatilities of Vodaphone PLC stock options as of 25 July 2003. The options traded on Vodaphone shares have maturities ranging from one month to two years. The x-axis is the "moneyness," defined as S/Xe^{-rT}, where S is the Vodaphone share price on 25 July 2003, X is the strike price, e is the base of the natural logarithm, r is the T-bill rate, and T is the option maturity. If Black–Scholes is correct, there can be only one value for implied volatility for all options of the same maturity. In Figure 4, the implied volatility at the low strike price is higher than that at the high strike price, and the difference is most marked for the short-maturity option.

Figure 4. Black–Scholes Implied Volatility for Vodaphone PLC Stock Options, 25 July 2003



Note: ATM means at the money.

If we try to fit a nonparametric risk-neutral density, $\hat{f}(S_T)$, such that prices of all European call options of a particular maturity T satisfy the following relationship,

$$c_j = e^{-rT}\int_{X_j}^{\infty}(S_T - X_j)\hat{f}(S_T)\,dS_T, \tag{7}$$

the fitted risk-neutral distribution will have large negative skewness and high kurtosis. On the one hand, the risk-neutral and actual stock price distributions do not have a strict one-to-one relationship (Cámara forthcoming), but we can at least conclude that the market does not price options based on the assumption that the stock price has a lognormal distribution or that stock returns have a normal distribution. Otherwise, the implied-volatility graph should be flat. On the other hand, as the time horizon increases, the distribution of long-horizon returns tends toward normal because of the central limit theorem. This conclusion is supported by the actual return data and the flatter implied volatility in Figure 4 for the options with the longer maturities.

Setting the Black–Scholes model aside, note that using the implied volatility of at-the-money (ATM) options is more popular in volatility forecasting than using the implied volatilities of the other options. The strong liquidity of ATM options also means that they are the least likely to be contaminated by pricing frictions. Implied volatility based on ATM options has been shown time and again (e.g., Christensen and Prabhala 1998; Fleming 1998; Ederington and Guan 2000; Li 2002) to have the greatest information content about future volatility, even if Black–Scholes is not the correct model for pricing options. Equation 5 is often used to correct any bias caused by model misspecification.

Thorny Outlier Issues

Outliers are large observations that come from a distribution different from the one generating day-to-day financial market variations. These outliers have a big impact on volatility estimation, modeling, and forecasting, but time-series volatility models based only on historical price information are ill designed for predicting unforeseen and unprecedented extreme events. Therefore, to penalize these *models for errors that arise because of unpredictable* outlier events is not logical. To reduce the influence of heavy tails and occasional large shocks, some have suggested that volatility modeling and forecast evaluation be based on absolute or logarithmic returns instead of squared returns (e.g., Pagan and Schwert 1990). The importance of tail events in financial markets and risk management cannot, however, be denied. So, outliers might be better studied separately with the use of a crisis model or techniques based on extreme-value theories.

If we accept the argument for separate evaluation, the next question is: How should one handle these outliers? The ways in which outliers have been tackled in the literature depend greatly on the outliers' size, the frequency of their occurrence, and whether the outliers produced an additive or a multiplicative impact.

For rare and additive outliers, the most common treatment is simply to remove them from the sample or omit them in the likelihood calculation (Kearns and Pagan 1993). For rare and multiplicative outliers that produce a residual impact on volatility, some researchers have included a dummy variable in the conditional volatility equation after the outlier returns have been dummied out in the mean equation (Blair, Poon, and Taylor 2001), as follows:

$$r_t = \mu + \psi_1 D_t + \varepsilon_t, \tag{8a}$$

with

$$\varepsilon_t = \sqrt{h_t}z_t \tag{8b}$$

and

$$h_t = \omega + \beta h_{t-1} + \alpha\varepsilon_{t-1}^2 + \psi_2 D_{t-1}, \tag{8c}$$

where ψ_1 and ψ_2 represent the crash impact on, respectively, the return and the conditional variance. In Blair et al., D_t is 1 when t refers to, for example, 19 October 1987 and 0 otherwise.

For outliers that occur more often, researchers may consider that the market has gone into a different mode and they may use a switching model (Friedman and Laibson 1989):[6]

$$r_t = \mu + \varepsilon_t, \tag{9a}$$

with

$$\varepsilon_t = \sqrt{h_t}z_t \tag{9b}$$

and

$$h_t = \omega + \beta h_t + \alpha F(\varepsilon_{t-1}^2), \tag{9c}$$

where F is calculated as

$$F(\varepsilon_{t-1}^2) = \begin{cases} \sin(a\varepsilon_{t-1}^2) & \text{if } a\varepsilon_{t-1}^2 < \dfrac{\pi}{2} \\ 1 & \text{if } a\varepsilon_{t-1}^2 \geq \dfrac{\pi}{2} \end{cases} \tag{9d}$$

and a is a constant term.

Researchers have documented that volatility caused by large returns (positive or negative) is less persistent than day-to-day volatility (Ederington and Lee 2001). If the outliers or group of adjacent large numbers are caused by a shift in volatility level, then such a level shift should be adjusted as in Aggarwal, Inclan, and Leal (1999):

$$r_t = \mu + \varepsilon_t, \tag{10a}$$

with

$$\varepsilon_t = \sqrt{h_t}z_t \tag{10b}$$

and

$$h_t = \omega + \beta h_{t-1} + d_1 D_1 + \ldots + d_n D_n + \alpha\varepsilon_{t-1}^2, \tag{10c}$$

where $D_1, \ldots, D_n$ are dummy variables taking a value of 1 from each point of sudden change of variance onwards and 0 otherwise.

The biggest difficulty in practice is that, even long after the outlier events, it is hard to identify which of these four cases the outlier belongs to—whether the event to be modeled is important because of size, frequency, additive impact, or multiplicative impact.

Option-implied volatility is a market-based volatility estimate and is the method least influenced by historical outliers, unless the outlier events fundamentally changed the option market's perception of future volatility. For example, some have claimed that the option market behaved as if it had "crashophobia" after the October 1987 market drop (Rubinstein 1994).

The SV models have a noise term in the volatility dynamic and are thus more flexible and less affected by large outliers than the ARCH models, which are, in turn, less severely affected than historical methods. Historical standard deviation will be affected by an outlier as long as it is in the volatility estimation period. For volatility estimation in all time-series models, we recommend trimming the outliers by imposing a cap on the largest values (see Huber 1981 for details) if one believes that the outlier event is an exception and not likely to be repeated.

Tips for Volatility Forecasters

All forecasting exercises consist of three main stages: Define the objectives of the forecast, develop and test competing models, and forecast the volatility values. All three stages involve complex issues, but the first stage crucially determines the course of action to be taken in the second and third stages. Here is some practical advice.

Stating the Objectives of Volatility Forecasting. First, be very clear about the objective (see the section "Objectives of Volatility Forecasting"), and accept the fact that no single model will fit all purposes. In risk management, for example, models for the tail distribution are needed.

Second, recognize what is being forecasted and its use. For example, if the volatility defined in a volatility swap contact is the standard deviation of a specified period, then you must adjust for option-implied bias. If the objective is to price an option, you must *not* correct for the implied volatility bias because the bias will be canceled out when implied volatility is fed back into the pricing model.

Building Volatility Models and Producing Forecasts. High-frequency data produce more accurate estimates for actual volatility and provide more accurate volatility forecasts than low-frequency data. Note, however, that the frequency should not be "ultrahigh." In a developed market, such as the United States, a five-minute interval has been generally recommended. The measurement interval will be longer for less liquid markets. Andersen and Bollerslev (1998) and Oomen (2004) provide some guidelines for determining the optimal frequency.

Volatility is a measure of average deviation from the mean. For a small sample, the sample mean is an extremely noisy estimate of the true mean in many financial time series. This flaw will have a direct impact on any volatility estimate or forecast. The mean estimate can be improved only by lengthening the sample period, not by sampling the data more frequently. Hence, a common practice in the stock and currency markets is to take deviation from zero based on the observation that the daily and weekly mean returns in speculative markets are close to zero.

Returns on speculative assets are not independently and identically distributed. Hence, variance of long-horizon returns is the aggregation (not the multiple) of single-period variances. The option-implied model provides volatility forecasts over the option's life. Any attempt to scale option-implied volatility to match a different horizon by using the square root of time will introduce error, the magnitude of which will depend on the slope of the volatility term structure.

Historical standard deviations are model free but greatly depend on how they are calculated (whether they are calculated from daily or weekly returns, whether the sample period is, for example, three or five years, whether the calculation covers overlapping periods, and so on). Conditional volatility models, such as ARCH and SV, and option-implied volatility models are spared these complications, but they are subject to model misspecifications.

Implied volatility for equity series is known to be unstable and is plagued with measurement errors and the variations caused by bid–ask spreads. Some intertemporal averaging (using, for example, the five-day average) and the use of past implied volatility as an instrumental variable have been shown to be helpful. Implied volatility usually dominates other volatility forecasts, but using the implied volatility of index options for the smaller markets, such as Sweden, works less well (Frennberg and Hansson 1995).

Option-implied volatility is also widely documented to be biased. It underforecasts low volatility and overforecasts high volatility; on average, implied-volatility estimates are greater than actual volatility. Because measurement error in option prices and noise in estimating actual volatility do not give a direction to the bias, the upwardly biased implied-volatility estimate has been linked to a volatility risk premium. Equation 5 provides an effective way to correct this bias.

Evaluating Volatility-Forecasting Methods. Be cautious about claims of superior forecasting performance. Take care to check that the study included out-of-sample forecasts and that the forecast error statistics differed significantly among models. What were the forecast evaluation criteria? If the evaluation was based on squared variance errors, then the standard error of the error statistics (often not reported) will be large because of the difficulty in estimating the fourth moment for thick tails.

Different cost functions will favor different forecasting methods. For example, nonlinear GARCH forecasts may produce smaller mean absolute errors than exponentially weighted moving average (EWMA) forecasts, but the tighter GARCH forecasts are likely to produce more VAR violations than EWMA forecasts.

As the forecast horizon lengthens, the advantage of sophisticated volatility models diminishes. For a horizon exceeding one year, Figlewski (1997) found that volatility forecasts derived from using low-frequency data from a sample period at least as long as the forecast horizon in the simple historical method produced the best result. Alford and Boatsman (1995) found that using median historical volatility of comparable companies adjusted for industry and size worked best for five-year-ahead equity volatility forecasts.

Conclusion

Financial market volatility is clearly forecastable. Research has shown that the forecasting power for stock index volatility is 50–58 percent for horizons of 1 to 20 trading days. The one-day-ahead forecasting record for exchange rates is 10–15 percent and is likely to increase by about threefold if the *ex post* volatility is measured more accurately. The one-week-ahead and one-month-ahead records for forecasting short-term interest rates have been documented to be, respectively, 8 percent and 24 percent. The current debate focuses on how far ahead one can accurately forecast and to what extent volatility changes can be predicted.

Based on the forecasting results, option-implied volatility dominates time-series models because the market option price fully incorporates current information and future volatility expectations. Between historical volatility and ARCH models, we found no clear winner, but they are both better than the stochastic volatility model. Despite the added flexibility and complexity of SV models, we found no clear evidence that they provide superior volatility forecasts. Also, high-frequency data clearly provide more information and produce better volatility forecasts, particularly over short horizons.

The conclusion that the option-implied method provides the best forecast does not violate market efficiency because accurate volatility forecasts do not conflict with underlying asset and option prices being correct.

Options are not available for all assets, so using historical volatility must be considered. These models are not necessarily less sophisticated than ARCH models. For example, the realized-volatility model of Andersen, Bollerslev, Diebold, and Labys (2003) is classified as a historical volatility model. The important aspects of using historical models are (1) that actual volatility be measured accurately and (2) that when high-frequency data are available, that information improves volatility estimation and forecasts.

A potentially useful area for future research is whether forecasting power can be enhanced by using exogenous variables. For example, Bittlingmayer (1998) linked volatility to macroeconomic news and systemwide factors; Spiro (1990) and Glosten et al. found a positive relationship between interest rates and volatility; Bollerslev and Jubinski (1999) found a positive relationship between trading volume and volatility; Hamilton and Lin (1996) showed that volatility is higher during recessions. Taylor and Xu (1997) fit 120 seasonal factors (representing hour, day, and week) to the conditional variance. What the literature has not yet shown is how these relationships improve volatility forecasts.

Ser-Huang Poon worked on this project while she was a visiting scholar in the Economics Department of the University of California at San Diego. She is grateful to the University of California at San Diego for its support.

Notes

1. In the early part of the sample period, we measured intraday returns at 30-minute and 15-minute intervals because the return series contained significant autocorrelations, possibly as a result of the less frequently traded stocks. In the more recent part of the sample period, 5-minute returns were used.
2. The autocorrelation coefficient measures the unconditional correlation between two series, whereas the partial autocorrelation coefficient measures the relationship between two series conditional on the relationships of all previous lags. For example, one would compute the partial for lag 2 by estimating the regression twice. The first regression would be the regression of the series on its lagged 1 values. The residual value of the first regression would then be used to regress on the series' lagged 2 values. The regression coefficient of the second regression would be the partial autocorrelation at lag 2.
3. The Federal Reserve's objective for open-market operations—purchases and sales of U.S. Treasury and federal agency securities—during the 1980s gradually shifted toward attaining a specified level of the federal funds rate.
4. The FI volatility models used in many papers allow a linear trend in volatility. One exception is the specification used by Bollerslev and Mikkelsen (1999). Hwang and Satchell (1998) made an adjustment specifically to remove this linear trend.
5. Extreme-value theory is a branch of statistics that has its main focus on the tail distribution. Returns and other observations that fall in the tail region are by definition large in magnitude and rare in occurrence.
6. Technically, frequent large numbers should not be called "outliers" because outliers should be rare.

References

Aggarwal, R., C. Inclan, and R. Leal. 1999. "Volatility in Emerging Stock Markets." *Journal of Financial and Quantitative Analysis*, vol. 34, no. 1 (March):33–55.

Alford, A.W., and J.R. Boatsman. 1995. "Predicting Long-Term Stock Return Volatility: Implications for Accounting and Valuation of Equity Derivatives." *Accounting Review*, vol. 70, no. 4 (October):599–618.

Andersen, T.G., and T. Bollerslev. 1998. "Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts." *International Economic Review*, vol. 39, no. 4 (November):885–905.

Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, and Paul Labys. 2003. "Modeling and Forecasting Realized Volatility." *Econometrica*, vol. 71, no. 2 (March):529–626.

Bittlingmayer, C. 1998. "Output, Stock Volatility, and Political Uncertainty in a Natural Experiment: Germany, 1880–1940." *Journal of Finance*, vol. 53, no. 6 (December):2243–57.

Black, F., and M. Scholes. 1973. "The Pricing of Options and Corporate Liabilities." *Journal of Political Economy*, vol. 81, no. 3 (May/June):637–654.

Blair, B., S. Poon, and S.J. Taylor. 2001. "Forecasting S&P 100 Volatility: The Incremental Information Content of Implied Volatilities and High-Frequency Index Returns." *Journal of Econometrics*, vol. 105, no. 1 (November):5–26.

Bollerslev, Tim, and D. Jubinski. 1999. "Equity Trading Volume and Volatility: Latent Information Arrivals and Common Long-Run Dependencies." *Journal of Business and Economic Statistics*, vol. 17, no. 1 (January):9–21.

Bollerslev, Tim, and Hans Ole Mikkelsen. 1999. "Long-Term Equity Anticipation Securities and Stock Market Volatility Dynamics." *Journal of Econometrics*, vol. 92, no. 1 (September): 75–99.

Cámara, António. Forthcoming. "Option Prices Sustained by Risk Preferences." *Journal of Business*.

Christensen, B.J., and N.R. Prabhala. 1998. "The Relation between Implied and Realized Volatility." *Journal of Financial Economics*, vol. 50, no. 2 (November):125–150.

Diebold, F.X., and A. Inoue. 2001. "Long Memory and Regime Switching." *Journal of Econometrics*, vol. 105, no. 1 (November):131–159.

Ederington, L.H., and W. Guan. 2000. "Measuring Implied Volatility: Is an Average Better?" Working paper, University of Oklahoma.

Ederington, L.H., and J.H. Lee. 2001. "Intraday Volatility in Interest-Rate and Foreign-Exchange Markets: ARCH, Announcement, and Seasonality Effects." *Journal of Futures Markets*, vol. 21, no. 6 (June):517–552.

Engle, R.F., and G.J. Lee. 1999. "A Long-Run and Short-Run Component Model of Stock Return Volatility." In *Cointegration, Causality, and Forecasting: A Festschrift in Honour of Clive W.J. Granger*. Edited by R. Engle and H. White. Oxford, U.K.: Oxford University Press.

Fama, Eugene F. 1965. "The Behavior of Stock-Market Prices." *Journal of Business*, vol. 38, no. 1 (January):34–105.

Figlewski, S. 1997. "Forecasting Volatility" *Financial Markets, Institutions and Instruments*, vol. 6, no. 1 (February):1–88.

Fleming, J. 1998. "The Quality of Market Volatility Forecasts Implied by S&P 100 Index Option Prices." *Journal of Empirical Finance*, vol. 5, no. 4 (October): 317–345.

Frennberg, P., and B. Hansson. 1995. "An Evaluation of Alternative Models for Predicting Stock Volatility: Evidence from a Small Stock Market." *Journal of International Financial Markets, Institutions and Money*, vol. 5, nos. 2/3:117–134.

Friedman, B.M., and D.I. Laibson. 1989. "Economic Implications of Extraordinary Movements in Stock Prices." *Brookings Papers on Economic Activity*, no. 2:137–172, 186–189.

Glosten, L.R., R. Jagannathan, and D.E. Runkle. 1993. "On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks." *Journal of Finance*, vol. 48, no. 5 (December):1779–1801.

Granger, C.W.J., and N. Hyung. 2004. "Occasional Structural Breaks and Long Memory with an Application to the S&P 500 Absolute Stock Returns." *Journal of Empirical Finance*, vol. 11, no. 3 (June):399–421.

Hamilton, J., and G. Lin. 1996. "Stock Market Volatility and the Business Cycle." *Journal of Applied Econometrics*, vol. 11, no. 5 (September/October):573–593.

Huber, Peter. 1981. *Robust Statistics*. New York: John Wiley & Sons.

Hwang, S., and S. Satchell. 1998. "Implied Volatility Forecasting: A Comparison of Different Procedures, Including Fractionally Integrated Models with Applications to UK Equity Options." In *Forecasting Volatility in the Financial Markets*. Edited by J. Knight and S. Satchell. London, U.K.: Butterworth-Heinemann.

Kearns, P., and A.R. Pagan. 1993. "Australian Stock Market Volatility: 1875–1987." *Economic Record*, vol. 69, no. 205 (June):163–178.

Latane, H., and R.J. Rendleman. 1976. "Standard Deviations of Stock Price Ratios Implied in Option Prices." *Journal of Finance*, vol. 31, no. 2 (May):369–381.

Li, K. 2002. "Long-Memory versus Option-Implied Volatility Prediction." *Journal of Derivatives*, vol. 9, no. 3 (Spring):9–25.

Mandelbrot, B. 1963. "The Variation of Certain Speculative Prices." *Journal of Business*, vol. 36, no. 4 (October):394–419.

Oomen, R.P. 2004. "Statistical Models for High Frequency Security Prices." Econometric Society 2004 North American Winter Meetings, no. 77: ideas.repec.org/p/ecm/nawm04/77.html.

Pagan, A.R., and G.W. Schwert. 1990. "Alternative Models for Conditional Stock Volatility." *Journal of Econometrics*, vol. 45, nos. 1–2 (July/August):267–290.

Poon, Ser-Huang, and Clive W.J. Granger. 2003. "Forecasting Volatility in Financial Markets: A Review." *Journal of Economic Literature*, vol. 41, no. 2 (June):478–539.

Rubinstein, M. 1994. "Implied Binomial Trees." *Journal of Finance*, vol. 49, no. 3 (July):771–818.

Spiro, Peter S. 1990. "The Impact of Interest Rate Changes on Stock Price Volatility." *Journal of Portfolio Management*, vol. 16, no. 2 (Winter):63–68.

Taylor, S.J. 1986. *Modelling Financial Time Series*. Chichester, U.K.: John Wiley & Sons.

Taylor, S.J., and X. Xu. 1997. "The Incremental Volatility Information in One Million Foreign Exchange Quotations." *Journal of Empirical Finance*, vol. 4, no. 4 (December):317–340.

Yoon, Gawon. 2003. "A Simple Model That Generates Stylized Facts of Returns." Working paper, Pusan National University.