A Source of Long Memory in Volatility

Namwon Hyung
University of Seoul

Ser-Huang Poon
University of Manchester

Clive W.J. Granger
University of California at San Diego

December 2006 (revised)

Abstract

This paper compares the out-of-sample forecasting performance of three long-memory volatility models (i.e., fractional integrated (FI), break and regime switching) against three short-memory models (i.e., GARCH, GJR and volatility component). Using S&P 500 returns, we find that structural break models produced the best out-of-sample forecasts, if future volatility breaks are known. Without knowing the future breaks, GJR models produced the best short-horizon forecasts and FI models dominated for volatility forecasts of 10 days and beyond. The results suggest that S&P 500 volatility is non-stationary at least in some time periods. Controlling for extreme events (e.g., the 1987 crash) significantly improved forecasting performance.

JEL Classifications: C22, C50, G10.

Keywords: Long memory; Structural breaks; Fractional integration; Volatility components; Regime switching; Volatility forecasting
1 Introduction

Financial market volatility is one of the most important attributes that affects the day-to-day operation of the Finance industry, as it is the key driver in investment analysis and risk management. Recently, there is an increase in trading activity on volatility as evidence by the volatility option contracts launched by the CBOE (Chicago Board of Option Exchange) in March 2006. Volatility in financial markets is known to cluster, a characteristic exploited in Engle’s (1982) ARCH model, which led to the model’s widespread use and Nobel Prize recognition in 2003. However, the length of volatility memory of ARCH and GARCH models is much shorter than what we often observe in long series of financial data. All volatility models differ in the volatility dynamics that they are designed to capture. A non-linear long-memory volatility model, for example, could produce a volatility forecast that is several percentage points different from a short-memory model, and each of these volatility percentage points is potentially worth up to $5,000 in a typical volatility swap contract. So there is a strong incentive to find out the true volatility dynamic and which model forecasts volatility more accurately.

Let $\rho_\tau$ denote the correlation between $x_t$ and $x_{t-\tau}$. A time series $x_t$ is said to have a short memory if $\sum_{\tau=1}^{n} \rho_\tau$ converges to a constant as $n$ becomes large. A series with long memory has autocorrelation values that decline slowly at a hyperbolic rate. The popular ARCH class of models is a short-memory volatility model. Granger and Joyeux (1980) and Hosking (1981) show that fractionally integrated series could produce the long-memory property described above. From these results, Ding, Granger and Engle (1993) propose a fractionally integrated volatility model based on absolute return $|r_t|^\delta$ where $\delta$ is a fraction. After the publication of Ding et. al, there has been a lot of research investigating if these fractionally integrated models could help to make better volatility forecasts and if long memory could explain anomalies in option prices.\footnote{For forecasting examples, see Li (2002), Vilasuso (2002), Andersen, Bollerslev, Diebold and Labys (2003), Pong, Shackleton, Taylor and Xu (2004), Martens and Zein (2004), Martens, van Dijk and de Pooter (2004), and Bhardwaj and Swanson (2006). For option pricing studies, see Bollerslev and Mikkelsen (1999) and Taylor (2000).}

Hitherto, most previous research has focused on linear models, fractionally integrated models in particular, to study long memory. More recent studies have shown that a number of non-linear volatility models can also produce long memory characteristics in volatility. Examples of such models include the break model (Granger and Hyung, 2004; Stărică and Granger,
2004) and the regime-switching model (Hamilton and Susmel, 1994; Diebold and Inoue, 2001; Hillebrand, 2005). In these two models, volatility is characterized by short memory between breaks and within each regime. Without controlling for the breaks and the changing regimes, volatility will produce spurious long-memory characteristics. Each of these models represents a very different volatility structure and produces volatility forecasts that are very different from each other and different from those of the fractionally integrated and short-memory models.

This paper analyzes the properties of the three long-memory volatility models (break, regime switching and FIGARCH), and compares their forecasting performance with classical short-memory models such as GARCH, GJR (Glosten Jogannathan and Runkle, 1993) and volatility component (Engle and Lee, 1999), and infinite-memory models such as exponential smoothing (ES) and random walk (RW). All of the contesting volatility models were fitted to the S&P 500 returns over a 29.5 year period from 4 January 1965 to 22 July 2003. Daily volatility forecasts were made using a rolling sample of 5,550 returns observations. The forecasting period covers 20.5 years from 21 January 1991 to 23 July 2003. Each day in the forecasting period, prediction horizons of various lengths were used (1, 5, 10, 20, 60 and 120 days). These horizons correspond to the horizons typically used in value-at-risk calculation, and they also match the maturity of exchange traded volatility contracts. Intraday returns were used to calculate the actual realized volatility for forecast evaluation.

Using daily returns of 155 financial series, we demonstrate that the so-called long memory in volatility is a marked phenomenon across different geographical regions and asset classes. With the use of intraday returns, we show that the S&P 500 realized volatility and its disentangled jump and diffusion components also display strong signs of long memory. Moreover, the S&P 500 volatility index (VIX), derived from options written on the S&P 500, also exhibits strong long-memory pattern. VIX is the “underlying” of VIX options and VIX futures traded on the CBOE. The long-memory feature will have an important implication for the pricing of these derivatives. To complement these empirical observations, we show by simulations that the volatility persistence of ARCH class models alone could not have accounted for this extraordinary long-memory pattern in financial market volatility.

The modelling and forecasting exercises in this paper show that the break model provides the best fit and forecast if future break dates and break sizes are known. Without the additional

---

2 Other non-linear examples, such as Franses, van der Leih and Paap (2002), Medeiros and Velga (2004) and Hillebrand and Medeiros (2006), can be viewed as a hybrid of break and regime-switching models.

3 The exponential smoothing model is similar to the exponentially weighted moving average (EWMA) model used by Riskmetrics. It can also be recast as an infinite-memory IGARCH model.
information and extra effort in forecasting breaks, GJR emerged as the best model for short-horizon volatility forecasts such as 1 and 5 days. For volatility forecasts of 10 days and beyond, the fractionally integrated model is the best, and at the very long horizons such as 60 and 120 days, its performance is similar to ES and RW. This result points toward an integrated model for long-horizon volatility forecasts.

Consistent with previous research (Blair, Poon and Taylor, 2001), the use of intraday data provides more accurate volatility estimate and better forecasting performance for all volatility models. However, the benefit of using intraday data is less important once the prediction horizon extends to 60 days and beyond. Moreover, the comparative ranking of volatility models does not change whether the intraday data or the daily squared return was used in calculating the actual volatility. This is a useful finding, as intraday data is not available for many financial time series and, more importantly, volatility and variance swaps are often defined based on returns calculated from daily closing prices and not intraday returns. We also find that controlling for the impact of outliers such as the 1987 stock market crash significantly reduces the mean absolute forecast errors, and in some cases significantly increases the correlation between the forecast and the actual volatility. Hence, to improve forecasting performance, it is important to handle these outliers carefully.

The rest of the paper is organized as follows: Section 2 presents evidence for long memory in financial market volatility. Section 3 presents the three long-memory volatility models, the short-memory and the integrated volatility models. Section 4 tests the in-sample fit and out-of-sample forecasting performance of these competing models. In Section 5, we analyse the break model in greater detail. Section 6 discusses the implication of breaks in volatility in a broader context and concludes.

2 How Much Volatility Memory is There?

Table 1 reports the sums of the first 1,000 autocorrelation coefficients for a selection of popular financial time series and the average values for a bigger collection covering 50 stock market indices, 25 equities, 25 exchange rates, 34 interest rates and 21 commodities. These are daily data collected from Datastream. They have different start dates (from 1965 to 1995) and the number of observations for each series ranges from 2,389 to 9,676. The autocorrelations are calculated for the absolute returns ($|r|$), squared returns ($r^2$), logarithm of absolute returns
The daily absolute return and squared return are often used as proxies for daily volatility and variance, respectively.

**TABLE 1 ABOUT HERE**

Outliers such as the 1987 stock market crash have an important impact on the summary statistics, and on the volatility forecasting result. Autocorrelation is the ratio of the autocovariance divided by the variance. The squaring of large values (such as that due to the 1987 stock market crash) creates a huge impact on variance calculation and a smaller impact on autocovariance. As a result, large values have the effect of dampening down the autocorrelation and weakening the detection of long memory in volatility proxies such as squared returns. Taking the absolute value reduces the impact of large values, making the long-memory driven autocorrelation effect more pronounced. Taking logarithms and trimming the data achieved the same result as taking absolute returns and in most cases make the long-memory effect of volatility even more prominent.

Table 1 also reports the differencing order, \( I(d) \), based on the Geweke and Porter-Hudak (1983) method. A long-memory process is covariance stationary if \( \sum_{\tau=1}^{n}(\rho_{\tau}/\tau^{2d-1}) \), for some positive \( d < 0.5 \), converges to a constant as \( n \to \infty \). When \( d \geq 0.5 \), the series is non-stationary. Based on the standard errors, not reported here, all but four \( d \) estimates are statistically and significantly greater than zero, and we cannot reject the null hypothesis that the differencing order is 0.5 for the majority cases. This means that most of these volatility proxies are extremely persistent and perhaps non-stationary also.

Table 2 reports the sum of the autocorrelation coefficients for the first 1,000 lags of the different components of the S&P500 realized volatility calculated from 10 minutes returns over the period 1 February 1983 to 31 July 2003 and the daily CBOE VIX, the implied volatility index derived from S&P 500 option prices, over the period 2 January 1990 to 31 July 2003. The realized volatility is separated into the continuous stochastic part and the jump part according to Barndorff-Nielsen and Shephard (2004; see the Appendix for details). The levels of persistence in all three realized volatility components and in the implied volatility index VIX are sizable and are much higher than the levels of persistence in most of the volatility

---

4For trimmed returns, we cap all values in the 0.1% tail to be the same as the 0.1% quantile. This is equivalent to trimming one daily return in every two years worth of data. Taking logarithms and trimming are efforts to curtail the impact of large values.

5The impact of outliers is relatively less severe on model estimations that are based on the maximisation of log-likelihood, as each observation only contributes to a small probability likelihood. The 1987 crash should not have too dramatic an impact on, e.g., GARCH parameter estimates, although if not controlled for, it is likely to produce very different forecasts, especially in the period immediately after the stock market crash.
proxies reported in Table 1.

TABLE 2 ABOUT HERE

The last panel of Table 2 reports the statistics for returns simulated from GARCH and GJR using parameters that represent high volatility persistence. The levels of volatility persistence generated by these models do not compare to all the empirical observations reported above and those reported in Table 1. The autocorrelation pattern of a large number of volatility series and different volatility measures for S&P 500 reported in this section show overwhelmingly that volatility is highly persistent. Whether this pronounced long-memory characteristic of volatility is real or spurious, and how best to model it, is the subject of the following sections.

3 Long-Memory Volatility Models

In this section, we present three models that are capable of producing characteristics of long memory in volatility such as those presented in Section 2. The first model, the fractionally integrated model presented in Section 3.1, is the simplest linear long-memory model. It is also the most commonly used and tested model in the literature. There are many other non-linear short-memory models that exhibit spurious long memory in volatility. Here, we study two of these: the break and regime-switching models. These two models plus the fractionally integrated model have very different volatility dynamics and produce very different volatility forecasts.

The volatility break model in Section 3.2 permits the mean level of volatility to change in a step function through time with some weak constraint on the number of breaks in volatility level. It is more general than the switching regime model in Section 3.3. In the case of the switching regime model, the mean levels of volatility differ according to the regimes, the total number of which is usually confined to a small number such as two or three.

In Section 2, we used the autocorrelation cumulant, $S_1 = \sum \rho_k$, as a measure of volatility persistence. Among the short-memory models, the sum of GARCH model parameters, $S_2 = \sum (\alpha_i + \beta_i)$, can also be used to measure volatility persistence, but this measure can only be used to compare volatility persistence among linear GARCH models. It is misleading to use $S_2$ to compare levels of volatility persistence across linear and non-linear models. For example, the break model can have a very large $S_1$ but, after adjusting for the break levels, a very small $S_2$. Moreover, for short-memory models, when $S_2 < 1$, $S_1$ is bounded. If $S_2 \geq 1$, $S_1$ is unbounded.
3.1 FI(E)GARCH Models

A fractionally integrated series has a differencing order between 0 and 1, whereas an integrated model has a differencing order of 1. The FIGARCH model has volatility persistence (measured either by $S_1$ or $S_2$) shorter than IGARCH but longer than GARCH. An IGARCH model is a GARCH model with the sum of persistence parameters, $S_2$, equal to 1. In this respect, the random walk and exponentially weighted moving average models can both be seen as variants of IGARCH. In the IGARCH model, a shock to the conditional variance has an impact on forecasts of the variance at all and infinite future horizons. In contrast, a fractional integrated series allows the impact of volatility shock to dissipate at a slow hyperbolic rate.

Granger (1980) and Andersen and Bollerslev (1997) suggest that cross-sectional aggregation of a large number of volatility components or news arrivals with different degrees of persistence could lead to long memory. In this case, the fractionally integrated model or an autoregressive model with a long distributed lag would be a natural model choice for such a data generating process. Baillie, Bollerslev and Mikkelsen (1996) report evidence of the presence of fractionally integrated behavior in the conditional variance of the nominal U.S. dollar-Deutschmark exchange rate. Bollerslev and Mikkelsen (1996) find that fractionally integrated models provide a better fit to S&P 500 returns than GARCH($p,q$) and IGARCH($p,q$) models and that the FIEGARCH specification is better than FIGARCH. Bollerslev and Mikkelsen (1999) find that FIEGARCH beats EGARCH and IEGARCH in pricing options of S&P500 LEAPS (Long-term Equity Anticipation Securities) contracts.

Specifically Bollerslev and Mikkelsen (1999) fitted an AR(2)-FIEGARCH(1, $d$, 1) as shown below:

\[
\begin{align*}
\ln h_t &= \omega_t + (1 + \psi_1 L) (1 - \phi_1 L)^{-1} (1 - L)^{-d} g (z_t), \\
g (z_t) &= \theta z_{t-1} + \gamma |z_{t-1} - E |z_{t-1}|, \\
\omega_t &= \omega + \ln (1 + \delta N_t),
\end{align*}
\]

where $\varepsilon_t = z_t h_t^{1/2}$ and $N_t$ is the number of non-trading days between day $t-1$ and $t$. The $-0.5h_t$ term in $r_t$ represents a convexity adjustment for lognormal variable. Note that EGARCH and FIEGARCH provide forecasts for $\ln h_t$. To infer a forecast for $h_t$ from $\ln h_t$ requires adjustment for Jensen’s inequality, which is not a straightforward task without the assumption of a normal distribution for $\ln h_t$; see Bollerslev and Mikkelsen (1999) and Taylor (2000). In this paper, we
use the FIGARCH$(1,d,1)$ model:

$$h_t = \omega + [1 - \beta_1 L - (1 - \phi_1 L)(1 - L)^d] \varepsilon_t^2 + \beta_1 h_{t-1}. \tag{2}$$

Unlike the FIEGARCH model, the FIGARCH model has an implicit upward trend in the conditional variance. To see this, one could rewrite the fractionally integrated component as an infinite sum of the residual term. Take the simplest FIGARCH$(0,d,0)$ case for example,

$$h_t = \omega + \left[1 - (1 - L)^d\right] \varepsilon_t^2.$$

For $d > -1$, the fractional difference operator $(1 - L)^d$ can be expressed as a Maclaurin series and gives

$$h_t = \omega + \varepsilon_t^2 - \sum_{j=0}^{\infty} \pi_j \varepsilon_{t-j}^2,$$

$$\pi_j = \frac{j - 1 - d}{j!}, \pi_0 = 1.$$

For $d < 0$, the sum $- \sum_{j=1}^{\infty} \pi_j$ does not converge. For $0 < d < 0.5$, the sum $- \sum_{j=1}^{\infty} \pi_j$ converges to 1. For the magnitude of $d$ estimated for S&P 500 and other financial series reported in Table 1, $h_t$, as a weighted sum of a long history of $\varepsilon_{t-j}^2$ whose value is very small (except during stock market crashes), converges as $j \to \infty$. The forecast value of $h_t$ is bounded, and the practice in the literature is to truncate the lag terms at $j = 1,000$. Thus, the implicit upward trend of $h_t$ projected from FIGARCH is not likely to be an issue here.

### 3.2 Breaks and Structural Change

Recent theoretical work on structural changes, switching regimes and occasional breaks has shown that any of these events is capable of producing the long-memory property.\textsuperscript{6} One interesting empirical finding comes from Aggarwal, Inclan and Leal (1999), who use the ICSS (Integrated Cumulative Sums of Squares) algorithm to identify sudden shifts in the variance of 20 stock market indices and the duration of such shifts. They find most volatility shifts are due to local political events. When dummy variables, indicating the location of sudden change in variance, were fitted to a GARCH$(1,1)$ model, most of the GARCH parameters became statistically insignificant.

\textsuperscript{6}See Diebold and Inoue (2001), Granger and Hyung (2004) and Mikosch and Stårică (2004) for examples. Engle and Smith (1999) examine a class of processes with random endogenous structural shifts at random intervals, and breaks of larger magnitude are classified as permanent shocks. Park (2002), on the other hand, models conditional heteroskedasticity as a non-linear function of integrated explanatory variables. In this paper, we do not study these two models.
Consider occasional shifts in the unconditional variance $\sigma_t^2$, which takes the value $\tau_i^2$ within each interval $i = 0, \ldots, R$. Let $k_i, i = 1, \ldots, R$ represent the break points and $1 < k_1 < k_2 < \ldots < k_R < T$.

$$\begin{align*}
\sigma_t^2 &= \tau_0^2 \quad 1 \leq t < k_1, \\
&\quad \cdots \\
&= \tau_R^2 \quad k_R \leq t < T.
\end{align*}$$

Granger and Hyung (2004) show that $d$, the integrating parameter of volatility, is a function of the number of break points, $R$, if $R$ is a bounded non-zero constant. When $d$ is bounded between 0 and 1, the expected value of $d$ is proportional to $R$.

To estimate the location of the break points, the number of breaks and the break sizes, we use a sequential search method. The sequential method was used by Bai (1997) and Bai and Perron (1998) for estimating multiple breaks in mean and by Inclan and Tiao (1994) for estimating breaks in variance. The sequential procedure starts with locating the first break point at $k$ where a hypothesis test of parameter constancy in the two subsamples $[1, k-1]$ and $[k, T]$ fails. Then the whole sample is divided into the two subsamples and parameter constancy tests are conducted for each of the subsamples. This procedure is repeated until the parameter constancy test is not rejected for all subsamples. The number of break points is equal to the number of subsamples minus 1.

Although asymptotic theory implies that the sequential procedure, coupled with hypothesis testing, produces a consistent estimate for the true number of breaks $R^*$, Bai (1997) shows by simulations that the procedure has a tendency to underestimate $R^*$. This problem can be overcome by using a two-step procedure as suggested by Bai (1997). In the first step, the goal is to obtain a consistent (or less biased) estimate for the error variance. This is achieved by allowing more breaks $R'$, solely for the purpose of constructing the error variance. It is evident that as long as $R' \geq R^*$, the error variance will be consistently estimated. The error variance is then used for estimating all the break points, either simultaneously using the Schwartz-Bayesian criterion as in Nunes, Kuan and Newbold (1995) or sequentially as described above.

Here, we use the ICSS algorithms from Inclan and Tiao (1994) to estimate the number of breaks and break positions in a GARCH(1,1)-with-breaks model as follows:

$$h_t = \omega_0 D_0 + \cdots + \omega_R D_R + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1},$$

where $D_0, \ldots, D_R$ are the $R+1$ dummy variables taking the value 1 in each regime of variance.
and zero elsewhere. We give this model an acronym B-GARCH(1,1).

3.3 Volatility Regime-Switching Model

The regime-switching ARCH model first appeared in Hamilton and Susmel (1994). It was extended to RS-GARCH by Gray (1996), Klaassen (2002), Bollen, Gray and Whaley (2000) and Liu (2000). Many of the successful applications appear to be in the modelling of interest rates that incorporate regime shifts and an interest rate level effect. Peria (2001) uses RS in economic panel data and allows the transition probabilities to vary according to economic conditions. Here, we use a 2-state RS-GARCH model:

\[ r_t \sim \begin{cases} 
N(\mu_1, h_{1t}) & \text{w.p. } p_{1t} \ \\
N(\mu_2, h_{2t}) & \text{w.p. } (1 - p_{1t}) \end{cases}, \]

\[ p_{1t} = \Pr(S_t = 1 | \Phi_{t-1}), \]

where \( h_t \) is integrated from the two states,

\[ h_t = E(r_t^2) - E(r_t)^2 = p_{1t}(\mu_1^2 + h_{1t}) + (1 - p_{1t})(\mu_2^2 + h_{2t}) - [p_{1t}\mu_1 + (1 - p_{1t})\mu_2]^2, \]

with transition probabilities \( \begin{pmatrix} P & 1 - P \\ 1 - Q & Q \end{pmatrix} \). In Gray (1996, p. 37) the switching probabilities \( P \) and \( Q \) are time-varying CDFs of the normal distribution such that \( P_t = \Phi(C_1 + d_1 r_{t-1}) \) and \( Q_t = \Phi(C_2 + d_2 r_{t-1}) \), whereas Peria (2001) uses the logistic function. Here we follow Gray (1996) with constant transition probabilities and use the normal CDF to specify the likelihood function,

\[ f_t(r_t) = \sum_{i=1}^{2} f_{it}(r_t) p_{it}, \]

where \( p_{it} \) is the probability of state \( i \) and

\[ f_{it}(r_t) = \frac{1}{\sqrt{2\pi h_{it}}} \exp \left\{ -\frac{(r_t - \mu_i)^2}{2h_{it}} \right\}. \]

Then by Bayes’ Rule, \( p_{1t} \) can be expressed in terms of \( p_{1t-1} \) by first writing

\[ g_{it} \equiv f_{it-1}(r_{t-1}) = \frac{1}{\sqrt{2\pi h_{it-1}}} \exp \left\{ -\frac{(r_{t-1} - \mu_i)^2}{2h_{it-1}} \right\}, \]

so that

\[ p_{1t} = (1 - Q) \frac{g_{2t-1}(1 - p_{1t-1})}{f_{1t-1} p_{1t-1} + f_{2t-1}(1 - p_{1t-1})} + P \frac{g_{1t-1} p_{1t-1}}{f_{1t-1} p_{1t-1} + f_{2t-1}(1 - p_{1t-1})}. \]
From the state probabilities, the log likelihood can be expressed as

\[ L = \sum_{t=1}^{T} \log [p_{1t} f_{1t} + (1 - p_{1t}) f_{2t}] . \]

4 Empirical Study

S&P 500 returns covering a 32.5-year period from 2 January 1969 to 23 July 2001 were used to test the volatility models for goodness of fit and for forecasting performance. In particular, two-thirds of the sample (22 years from 2 January 1969 to 18 January 1991) was used for estimation and one third of the sample (10.5 years from 21 January 1991 to 23 July 2001) was used for out-of-sample forecast evaluation. We have taken care to remove all the zero returns that are due to holidays. The returns were not adjusted for dividend distributions as dividend information is not available prior to 1983, and because of the fact that dividend distribution does not appear to alter the parameters estimates significantly when dividend adjustment was made for the post-1983 period.

4.1 Short-Memory Models

To serve as benchmarks for the long-memory models, we estimate three short-memory models: GARCH(1,1), GJR(1,1) and the volatility component model, VC(1,1). The GARCH(1,1) model is given by

\[ h_t = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}. \]

The GJR(1,1) model allows for asymmetric responses of conditional volatility to lagged squared returns and can be expressed as

\[ h_t = \omega + \alpha_1 \varepsilon_{t-1}^2 I_{[\varepsilon_{t-1} \leq 1]} + \gamma_1 \varepsilon_{t-1}^2 I_{[\varepsilon_{t-1} > 1]} + \beta_1 h_{t-1}, \]

where \( I_{[\cdot]} \) is an indicator function taking the value of 1 if the specified condition is met and zero otherwise. Engle and Lee (1999) model the volatility process as the sum of a permanent process, \( m_t \), that has memory close to a unit root, and a transitory mean reverting process, \( u_t \), that has a more rapid time decay. The VC(1,1) model can be written as a GARCH(2,2) with the conditional variance mean-reverting to a long term trend level, \( m_t \), instead of a fixed position at \( \sigma \). Specifically, \( m_t \) is permitted to evolve slowly in an autoregressive manner:

\[ (h_t - m_t) = \alpha (\varepsilon_{t-1}^2 - m_{t-1}) + \beta (h_{t-1} - m_{t-1}) \equiv u_t, \quad (3) \]

\[ m_t = \omega + \rho m_{t-1} + \varphi (\varepsilon_{t-1}^2 - h_{t-1}), \]
where \((h_t - m_t) = u_t\) represents the short-run transitory component. The reduced form of equation (3) can be expressed as a GARCH(2,2) process, as shown below:

\[
h_t = (1 - \alpha - \beta) \omega + (\alpha + \varphi) \varepsilon_{t-1}^2 + \left[-\varphi (\alpha + \beta) - \alpha \rho\right] \varepsilon_{t-2}^2 \\
+ (\rho + \beta - \varphi) h_{t-1} + [\varphi (\alpha + \beta) - \beta \rho] h_{t-2},
\]

where all five parameters, \(\alpha, \beta, \omega, \varphi\) and \(\rho\), are constrained to be positive and real; \(0 < (\alpha + \beta) < \rho < 1\) and \(0 < \varphi < \beta\).

The volatility component model has various interesting properties: (i) both \(m_t\) and \(u_t\) are driven by \((\varepsilon_{t-1}^2 - h_{t-1})\); (ii) the short-run volatility component mean-reverts to zero at a geometric rate of \((\alpha + \beta)\) if \(0 < (\alpha + \beta) < 1\); (iii) the long-run volatility component evolves over time following an AR process and converges to a constant level defined by \(\omega/(1 - \rho)\) if \(0 < \rho < 1\); (iv) it is assumed that \(0 < (\alpha + \beta) < \rho < 1\) so that the long-run component is more persistent than the short-run component.

The VC model was found to describe several economic and asset pricing relationships. Many have proposed that the volatility persistence of large jumps is shorter than shocks due to ordinary news events. The component model allows large shocks to be transitory. Indeed, Engle and Lee (1999) establish that the impact of the October 1987 crash on stock market volatility was temporary. The expected risk premium, as measured by the expected amount of returns in excess of the risk free interest rate, in the stock market was found to be related to the long-run component of stock return volatility.\(^7\) The well documented “leverage effect” (or volatility asymmetry) in the stock market is shown to have a temporary impact; the long-run volatility component shows no asymmetric response to market changes.\(^8\)

### 4.2 Estimation

All volatility models were fitted assuming the following specification for returns:

\[
\begin{align*}
r_t &= \mu + \varepsilon_t, \\
\varepsilon_t &= z_t \sqrt{h_t}, \quad z_t \sim N(0,1),
\end{align*}
\]

and estimated using quasi maximum likelihood. Inferences were made using Bollerslev and Wooldridge (1992)'s robust standard errors. We fixed all the pre-sample values of \(\varepsilon_t^2\) for

---

\(^7\)The relationship between risk premium and “total” volatility was first studied in Merton (1980) and French, Schwert and Stambaugh (1987).

\(^8\)The leverage effect was first documented by Black (1976) and Christie (1982) and subsequently modelled by Nelson (1991).
At the unconditional sample variance. The approach taken here for estimation of the FIGARCH model is based on a truncated ARCH representation. The FIGARCH model in (2) can be transformed to ARCH model with infinite lags as follows:

\[ h_t = \frac{\omega}{1 - \beta} + \left[ 1 - \frac{(1 - \phi L)(1 - L)^d}{(1 - \beta)} \right] \epsilon_t^2, \]

writing \( \beta_1 \equiv \beta \) and \( \phi_1 \equiv \phi \). Given that \((1 - L)^d = \sum_{j=0}^{\infty} \pi_j L^j \) and \( \pi_0 = 1 \), the parameters in the lag polynomials in \([\cdot]\) may be written as

\[
\lambda(L^j) = \begin{cases} 
\phi - \beta - \pi_1, & \text{for } j = 1, \\
\beta \lambda_{j-1} + (\pi_j - \phi \pi_{j-1}) & \text{for } j \geq 2.
\end{cases}
\]

Following Bollerslev and Mikkelsen (1996), the estimation was made by truncating all lags beyond \( j = 1,000 \).

The break points in the B-GARCH model (i.e., GARCH with occasional breaks in variance) was estimated using the ICSS algorithm from Inclan and Tiao (1994). The minimum duration between two consecutive breaks is set at 20 days to reduce the possibility that any temporary shocks is being mistaken as a break. Following Aggarwal, Inclan and Leal (1999), the estimation was done in two stages. The volatility change points were estimated in the first stage. In the second stage, dummy variables representing the different volatility regimes were introduced into the conditional variance equation of the GARCH model to account for the volatility level shifts.

\[ \text{TABLE 3 ABOUT HERE} \]

Table 3 presents parameter estimates, robust \( t \)-ratios and log-likelihood values. In terms of the log-likelihood, the B-GARCH models came out on top, principally due to the large number of break points estimated, 27 in total for the 22-year period. After the B-GARCH model, and in order of best fit, were RS-GARCH, GJR, FIGARCH, VC and GARCH. It is interesting to note that GJR, a short-memory model with leverage effect, produced a better fit than FIGARCH and VC, both with many more parameters. The goodness of fit of FIGARCH and VC were similar. The GARCH(1,1) model without the leverage and long-memory features produced the worst fit.

The parameter estimates reported in Table 3 also revealed sharp contrasts in volatility structure and persistence represented by the six volatility models. For the short-memory models, the volatility persistence is close to 1 (i.e., 0.991 for VC, 0.983 for GARCH and 0.982 for GJR). FIGARCH has a differencing parameter \( d \) that is is greater than 0.5, which
technically means the conditional variance is non-stationary. In contrast, the B-GARCH and the RS-GARCH models point towards a much smaller degree of volatility persistence once the volatility level shifts are accounted for. In the B-GARCH case, volatility persistence is almost non-existent. For the RS-GARCH, volatility is more persistent in the high volatility state, but at a level much smaller than those of FIGARCH, VC, GARCH and GJR.

**FIGURE 1 ABOUT HERE**

In general, the volatility break levels, permanent component and regimes in the different model estimates correspond well with each other. RS-GARCH models the conditional variance switching between high and low volatility regimes. From Table 3, the unconditional variance, $\omega_i/[1 - (\alpha_i + \beta_i)]$, is 0.456 and 2.682 for the two regimes (or 0.675 and 1.638 in terms of unconditional standard deviation). In Figure 1(a) to 1(c), we plot the conditional standard deviation estimated from the VC model and its two components, the permanent and transitory components. In Figure 1(d) and 1(e), we plot the volatility breaks estimated by two different methods: (i) the change-in-mean approach of Bai (1997) and applied to the intraday realized volatility; (ii) the change in variance approach in Inclan and Tiao (1994) and applied to S&P 500 returns. Figures 1(b), 1(d) and 1(e) share a common pattern. All these figures have sharp peaks or volatility breaks around 1974 (oil price crisis), 1987 (stock market crash) and post 1997 (Asian crisis). The permanent volatility in Figure 1(b) is jagged by design, whereas the two break models have volatility moving as a step function. The Bai (1997) method produces more breaks than the Inclan and Tiao (1994) method, but we should bear in mind that they are applied to different data sets pertaining to the S&P 500. So the different number of break points could be due to the different data sets or different methods or both.

### 4.3 Out-of-Sample Forecasts

Daily volatility forecasts were made using a rolling sample of 5,550 returns observations. The forecasting period covers 10.5 years from 23 January 1991 to 23 July 2001. Each day in the forecasting period, prediction horizons of various lengths were used: 1, 5, 10, 20, 60 and 120 days. These prediction horizons correspond to the durations used in value-at-risk calculation, portfolio revision and pricing option. Intraday data was used to calculate actual and realized volatility for forecast evaluation.
4.3.1 Recursive Procedure

In this section, we will use $\hat{h}$ to denote the forecast variance. When there is only one subscript, $\hat{h}_{t+j|t}$ refers to the point forecast for day $t+j$. The conditional operator means that the forecast $\hat{h}_{t+j|t}$ is made using all information available on day $t$. When there are two time subscripts, $\hat{h}_{t,s} = \sum_{i=1}^{s} \hat{h}_{t+i|t}$ refers to the forecast for variance over the period from $t+1$ to $t+s$ using all information available at time $t$.

To make a forecast for time $t+1$, we use model parameters estimated from the past 5,550 returns. The estimation period is then rolled forward one day, deleting the observation at time $t-5549$ and adding the observation at time $t+1$. The model is then re-estimated and a forecast is made for time $t+2$. This rolling method is repeated until the end of the out-of-sample forecast period. The one-step-ahead forecast procedure provides predictions for January 21, 1991 to July 23, 2001 inclusive with 2,652 daily point estimates. On each day, forecasts are also made for the next 5-, 10-, 20-, 60- and 120-day volatility.

The measure of daily realized variance is the daily squared intraday returns:

$$r_{(L),t}^2 \equiv \sum_{l=1}^{L} r_{t,l}^2,$$

where $L$ is the number of intraday intervals.\(^9\) So the $s$-period cumulated realized variance for the time $t+1$ to $t+s$ is

$$\sum_{i=1}^{s} r_{(L),t+i}^2, \quad s = 1, 5, 10, 20, 60, 120. \quad (4)$$

It is used for evaluating the $s$-period forecast $\hat{h}_{t,s} = \sum_{i=1}^{s} \hat{h}_{t+i|t}$.

As the prediction horizon increases, there is a huge amount of information overlap between two adjacent forecasts. Take the most extreme case of $s = 120$, for example, where 119 of the 120 days are the same for $\hat{h}_{t,s}$ and $\hat{h}_{t+1,s}$. Such a substantial data overlap will induce serial correlation in the forecasts and forecast errors. To control for potential bias and serial correlation, we use the Hansen and Hodrick (1980) and Diebold and Mariano (1995) procedures in calculating standard errors in all statistical tests below.

---

\(^9\)When realized volatility is derived from intraday data, there will be market microstructure noise (see Zhang, Mykland and Ait-Sahalia, 2005; Hansen and Lunde, 2006). Here, we use the autocorrelation of intraday returns to gauge the problem of microstructure noise. As a result, the intraday interval is chosen to be 10 minutes for the more recent sample period. For the earlier sample period, where there was less liquidity, the intraday interval used is 30 minutes. The overnight volatility is equal to squared close-to-open return.
4.3.2 One- and Multi-Step Forecasts

The s-step forecast $\hat{h}_{t+s|t}$ is defined by appropriate substitution based on the conditional volatility specification, so for the GARCH(1,1) model,

$$\hat{h}_{t+s|t} = \omega + \alpha_1 \varepsilon^2_{t+s-1|t} + \beta_1 \hat{h}_{t+s-1|t}.$$  

Using the fact that at time $t$, $\varepsilon^2_{t+i|t} = \varepsilon^2_{t+i}$ and $\hat{h}_{t+i|t} = h_{t+i}$ for $i \leq 0$, the one-step forecast is

$$\hat{h}_{t+1|t} = \omega + \alpha_1 \varepsilon^2_{t} + \beta_1 h_t.$$  

The forecast of multi-period cumulated volatility $\hat{h}_{t,s}$ is taken to be the sum of individual multi-step point forecasts $\sum_{i=1}^{s} \hat{h}_{t+i|t}$. These multi-step point forecasts can be produced by recursive substitutions using $\varepsilon^2_{t+i|t} = \hat{h}_{t+i|t}$ for $i > 0$:

$$\hat{h}_{t+s|t} = \omega + (\alpha_1 + \beta_1) \hat{h}_{t+s-1|t}, \text{ for } s > 1.$$  

For the GJR-GARCH(1,1) model,

$$\hat{h}_{t+1|t} = \omega + \alpha_1 \varepsilon^2_{t} I_{[\varepsilon_t \leq 1]} + \gamma_1 \varepsilon^2_{t} I_{[\varepsilon_t > 1]} + \beta_1 h_t,$$

$$\hat{h}_{t+s|t} = \omega + \left[ \frac{1}{2} (\alpha_1 + \gamma_1) + \beta_1 \right] \hat{h}_{t+s-1|t}, \text{ for } s > 1.$$  

For the FIGARCH(1,d,1) model,

$$\hat{h}_{t+s|t} = \frac{\omega}{1 - \beta_1} + \left[ 1 - \frac{(1 - \phi_1 L)(1 - L)^d}{1 - \beta_1 L} \right] \varepsilon^2_{t+s-1|t}.$$  

For the model with occasional breaks in variance, the multiple break points are estimated by the ICSS algorithm. The conditional variance series is then separated into regimes based on the identified break dates and the sample variance in each regime was calculated. If we do not have the ability to predict future breaks, then the forecasts were made assuming that there is no further break in the future. Hence, the unconditional variance for all future periods is the same as that at time $t$. From here, the construction of volatility forecasts follows that for GARCH(1,1) or GJR-GARCH(1,1) described above. Take the B-GARCH(1,1) case for example,

$$\hat{h}_{t+s|t} = \omega_{R+1} + \alpha_1 \varepsilon^2_{t+s-1|t} + \beta_1 \hat{h}_{t+s-1|t}.$$  

For the VC model, the multi-step forecast is produced from

$$\hat{h}_{t+s|t} = \hat{q}_{t+s|t} - (\alpha + \beta) \hat{q}_{t+s-1|t} + (\alpha + \beta) \hat{h}_{t+s-1|t},$$

$$\hat{q}_{t+s|t} = \omega + \rho \hat{q}_{t+s-1|t} + \varphi \left( \varepsilon^2_{t+s-1|t} - \hat{q}_{t+s-1|t} \right).$$
The volatility regime-switching model has a GARCH(1,1) specification in each volatility regime. So the individual RS-GARCH(1,1) forecast is the same as the GARCH(1,1) model and the forecast for future volatility integrated across the two regimes is calculated as

$$\hat{h}_{t+s|t} = p_{1,t+s|t} \left( \mu_1^2 + \hat{h}_{1,t+s|t} \right) + (1 - p_{1,t+s|t}) \left( \mu_2^2 + \hat{h}_{2,t+s|t} \right)$$

$$- \{ p_{1,t+s|t} \mu_1 + (1 - p_{1,t+s|t}) \mu_2 \}^2.$$

The forecast for conditional variance, $\hat{h}_{i,t+s|t}$ for $i = 1, 2$, is constructed in a similar fashion as the GARCH(1,1) model described above.

For the purpose of comparison, volatility forecasts are also produced for exponential smoothing and random walk models. For the exponential smoothing model,

$$h_t = (1 - \beta) h_{t-1} + \beta \gamma_{t-1}^2 + \varepsilon_t \quad \text{and} \quad 0 \leq \beta \leq 1,$$

where the smoothing parameter $\beta$ is estimated by minimizing the in-sample forecast errors $\varepsilon_t$. The forecasts from the exponential smoothing model are

$$\hat{h}_{t+1|t} = (1 - \beta) h_t + \beta \gamma_t^2,$$

$$\hat{h}_{t+s|t} = \hat{h}_{t+1|t}, \text{ for } s > 1.$$

The forecast for the random walk model is simply

$$\hat{h}_{t+i|t} = \hat{h}_{t+i-1|t} = \cdots = h_t,$$

for all future periods. For both the exponential smoothing and the random walk models, we use squared daily return to proxy for the realized variance.

### 4.3.3 Simulated Forecasts

Before we delve into the empirical estimation and forecasting results, it is useful to have good intuition about the dynamics of volatility forecasts generated by the six volatility models. First, we simulated six sets of returns generated using $N(0, 1)$ random variates together with the parameter values reported in Table 3. Then we calculate the autocorrelation coefficients of the absolute returns for the first 100 lags and produce volatility point forecasts for the first 120 days using the corresponding volatility model.

Figure 2 presents the autocorrelation coefficients and volatility point forecasts for the six volatility models. Corresponding to the theoretical predictions, the autocorrelation coefficients...
of absolute returns decay fairly quickly for the GARCH and GJR-GARCH models. The RS-GARCH and VC models have a longer distributed lags of positive autocorrelations, whereas the autocorrelation coefficients of absolute returns generated from the FIGARCH and B-GARCH models hover above zero for up to 100 lags with extremely slow decay.

**FIGURE 2 ABOUT HERE**

In terms of volatility forecasts, the point forecast converges to unconditional variance very quickly for the B-GARCH and RS-GARCH models because the \((\alpha_i + \beta_i)\) values for these two models are small once the volatility breaks or regime shifts are separately accounted for. The reverse is true for the GARCH, GJR-GARCH and VC models for which the \((\alpha_i + \beta_i)\) values are very high. Full convergence to the unconditional variance did not take place for these three models even after 120 days. The FIGARCH model produced a pattern in between the two extremes.

These simulations highlight the crucial fact that the volatility forecast is largely dominated by the volatility persistence parameters, \(\alpha_i\) and \(\beta_i\), of the conditional variance equation and the projected future level of the unconditional variance. In these simulations, we assume that there are no future volatility breaks or regime shifts, or at least the modeler is not able to predict these. Even with this simplifying assumption, the volatility forecasts from the six models are startlingly different.

**4.4 Forecast Evaluation**

For forecast evaluation, we use the mean absolute forecast error (MAE) criteria and the regression based forecast efficiency and orthogonality test. The MAE statistic for a set of \(m\) \(s\)-step forecasts is computed as

\[
MAE = \frac{1}{m} \sum_{t=0}^{m-1} |y_{t,s} - x_{t,s}|
\]

where \(y_{t,s} \equiv h_{t,s} = \sum_{i=1}^{s} h_{t+i}, x_{t,s} \equiv \hat{h}_{t,s} = \sum_{i=1}^{s} \hat{h}_{t+i,t}\). The null hypothesis of equality of forecast performance from different models are checked using the Diebold and Mariano (DM, 1995) test procedure. First define the loss differential between the absolute forecast errors from two models as

\[
d_t \equiv |e_{1t}| - |e_{2t}|.
\]
The null hypothesis of equal forecast accuracy implies zero expected loss differential \( E (d_t) = 0 \). The DM test targets the mean, 
\[
\bar{d} = \frac{1}{m} \sum_{t=1}^{m} d_t,
\]
with the test statistic, 
\[
S_1 = \frac{\bar{d}}{\sqrt{\frac{1}{m} 2\pi \hat{f}_d (0)}}.
\]
\[
2\pi \hat{f}_d (0) = \sum_{\tau=-(m-1)}^{m-1} 1 \left( \frac{\tau}{S(m)} \right) \hat{\gamma}_d (\tau),
\]
\[
\hat{\gamma}_d (\tau) = \frac{1}{m} \sum_{i=|\tau|+1}^{m} (d_t - \bar{d}) (d_{t-|\tau|} - \bar{d}).
\]
The operator \( 1(\tau/S(m)) \) represents the lag window, and \( S(m) \) is the truncation lag with
\[
1 \left( \frac{\tau}{S(m)} \right) = \begin{cases} 
1 & \text{for } \left| \frac{\tau}{S(m)} \right| \leq 1, \\
0 & \text{otherwise}.
\end{cases}
\]
Assuming that \( s \)-step ahead forecast errors are at most \( (s-1) \)-dependent, Diebold and Mariano (1995) recommend using \( S(m) = s - 1 \). It is not likely that \( \hat{f}_d (0) \) will be negative, but in the rare event that \( \hat{f}_d (0) < 0 \), it should be treated as zero and the null hypothesis of equal forecast accuracy be rejected automatically. Diebold and Mariano (1995) show that \( S_1 \sim N(0,1) \) under the null hypothesis of zero expected loss differential.

To check forecast efficiency, we regress the forecasts on the actual volatility, as shown below:
\[
y_{t,s} = \alpha + \beta x_{t,s} + \nu_t,
\]
where \( x_{t,s} \) is the square root of the cumulated sum of \( s \)-period variance forecast and \( y_{t,s} \) is the square root of the cumulated sum of the actual realized volatility for the date \( t + 1 \) to \( t + s \). The prediction is unbiased if and only if \( \alpha = 0 \) and \( \beta = 1 \).

However as Stărică and Granger (2005) point out, any possibility of non-stationarity of volatility will make the regression test in (5) invalid. To take into account such a possibility, we have also tested if \( \alpha = 0 \) and \( \beta = 1 \) in the differenced version of regression (5),
\[
\tilde{y}_{t,s} = \alpha + \beta \tilde{x}_{t,s} + \nu_t,
\]
where \( \tilde{y}_{t,s} = y_{t,s} - \sqrt{sh_t}, \) \( \tilde{x}_{t,s} = x_{t,s} - \sqrt{sh_t} \).

Since the error term, \( \nu_t \), is heteroskedastic and serially correlated when overlapping data are used, standard errors of the parameter estimates are computed based on Hansen and Hodrick
Let $X$ be the row matrix of regressors including the constant term. Then

$$
\hat{\Psi} = m^{-1} \sum_{t=1}^{m} \nu_{t}^{2} X_{t}'X_{t} + m^{-1} \sum_{k=1}^{m} \sum_{t=k+1}^{m} Q(k,t) \nu_{k} \nu_{t} \left( X_{k}'X_{k} + X_{t}'X_{t} \right),
$$

where $\nu_{k}$ and $\nu_{t}$ are the residuals for observation $k$ and $t$ from the regression. The operator $Q(k,t)$ is an indicator function taking the value 1 if there is information overlap between $X_{k}$ and $X_{t}$. The adjusted covariance matrix for the regression coefficients is then calculated as

$$
\hat{\Omega} = \left( X'X \right)^{-1} \hat{\Psi} \left( X'X \right)^{-1}.
$$

Canina and Figlewski (1993) conduct some simulation studies and find that the corrected standard errors in (6) are close to the true values. They support the use of overlapping data as it reduces the standard error between one-quarter to one-eighth of what would be obtained with non-overlapping data.

4.5 Results

Table 4 reports forecasting results for the six volatility models described previously together with the results for exponential smoothing (ES) and random walk (RW) models, both of which could be described as integrated models with infinite volatility memory. The actual volatility is calculated using intraday return data as described in Section 4.3.1. Using R-square in the last column of Table 4 as the performance criteria, GJR, FIGARCH and RW emerge in this order as the best volatility forecasting models as the prediction horizon increases. The explanatory power is good for all models, and the R-square measures range from 35–50% for 1- to 120-day volatility forecasts, with the predictive power increasing as the forecast horizon increases.

**TABLE 4** ABOUT HERE

Table 4 also reports the MAE (mean absolute forecast error) and tests of significant differences in forecast accuracy using the Diebold and Mariano test statistic described in Section 4.4. For each forecast horizon, the benchmark is the model with the smallest MAE for that particular forecast horizon. We then check forecast supremacy of the best model by comparing the benchmark with all the other models using the Diebold and Mariano test. Cases where the null hypothesis of equal forecast accuracy was rejected are marked with “*”. Based on the MAE criterion, GJR, FIGARCH, ES and RW emerge in this order as the best volatility forecasting models as prediction horizon increases from 1 to 120 days.

The forecast biasedness test produced similar outcomes for all volatility models with the exception of RS-GARCH. In the majority of the cases, $\alpha > 0$ and $\beta < 1$, which means that all
volatility models, except RS-GARCH, under-forecast in high volatility states and over-forecast in low volatility states. Here, we do not distinguish the sign and the different impacts of mis-forecast in the high and low volatility states. Indeed, should we consider the economic costs of mis-forecasts, the ranking of the models could well be very different.

Table 5 produces the same analysis as Table 4 but uses the daily squared returns, instead of the intraday returns, to proxy for actual volatility. Consistent with Blair, Poon and Taylor (2001), the smaller R-square measures in Table 5 compared with those in Table 4 indicate that daily squared return is a very noisy proxy for actual volatility. The R-square of all volatility models was reduced by 25% on average for the 1-day forecast and 15% on average for the 5-day forecast. For 10-day and 20-day forecasts, the R-square reductions are smaller (9% and 3%, respectively). The benefit of using intraday data becomes less important once the prediction horizon extends to 60 days and beyond. Results in Table 5 shows that for forecasting 60- to 120-day volatility, the two estimates give similar explanatory power. Interestingly, the comparative ranking of volatility models does not change whether the intraday data or the daily squared returns were used in calculating actual volatility. This is a useful finding, as we can now study the usefulness of volatility forecasting models on a wider set of financial series where intraday data are not available. Although detail results are not reported here, we find that controlling for the 1987 stock market crash with dummy variables significantly reduced the mean absolute forecast errors and in some cases significantly increased the R-square.

**TABLE 5 ABOUT HERE**

Table 6 repeats the analysis in Table 4 but performs the regression tests using volatility differences instead of volatility levels, to take into account the possibility that volatility may be non-stationary. The MAE values are the same as those reported in Table 4, since the MAE definition has not changed. The most important change in the results is the R-square. The R-square of the 1-day forecast dropped by 10% on average across all volatility models, and for the longest forecast horizon (120-day forecast), the R-square rose by 20% on average across all models. The drop in R-square for the 1-day forecast is a sign that volatility over short intervals could be non-stationary. The increase in R-square for the long-horizon forecasts highlights the potential of these volatility models in detecting changes in volatility amongst long-term volatility trends.

**TABLE 6 ABOUT HERE**
4.5.1 Summary

In summary, the results presented in this section show that, despite a relatively poor fit, the fractionally integrated model emerged as the best model for long-horizon volatility forecasts. The GJR model provides a relatively good fit and dominates all other volatility models for short-horizon forecasts. At long horizons, GJR has a smaller R-square and a bigger MAE than FIGARCH, but the null hypothesis of equal forecast accuracy between GJR and FIGARCH was not rejected statistically.

All the other key findings can be summarized as follows:

(i) Tables 1 and 2 demonstrate that long memory in volatility is an overwhelming phenomenon and is inconsistent with the characteristics of popular short-memory models such as GARCH and GJR.

(ii) According to the in-sample estimation results reported in Table 3, the models ranked in the order of best fit are B-GARCH, GJR, RS-GARCH, FIGARCH, VC and GARCH. This set of results signals the importance of volatility level shifts and asymmetry in the modelling process.

(iii) Figure 1 shows that the mean level of volatility fluctuates roughly between high and low volatility regimes, but does not remain fixed within the two states. The RS-GARCH model fitted here assumes that the unconditional variance is constant within each of the two volatility regimes, and hence constant for the entire series (given the probability of occurrence of each regime). For this reason, B-GARCH provides a much better in-sample fit and out-of-sample forecasts than RS-GARCH, as B-GARCH places no constraints at all on the unconditional variance. From here onwards, we will concentrate our analyses on B-GARCH only and drop RS-GARCH from further investigation.

(iv) The simulation results presented in Figure 2 show that the parameters estimated from real data and reported in Table 3 produced very high volatility persistence for GARCH, GJR and VC, and very low volatility persistence for B-GARCH and RS-GARCH. Volatility persistence is significantly reduced once volatility breaks and regime shifts are accounted for.

(v) Without incorporating volatility breaks in the forecasting period, B-GARCH is more likely to under-forecast because of the low volatility persistence parameters. This is reflected
in the low $\beta$ coefficient for B-GARCH in the regression efficiency test. The $\beta$ coefficient for B-GARCH is the lowest among the six volatility models.

5 The Case for Volatility Breaks

In the previous section, contradictory results were obtained between estimation and prediction. In this section, we study the model properties and provide some explanations for this puzzling result. Here, we argue that volatility breaks and changes in unconditional variance is the likely cause of the widely observed long memory in volatility. We will show in Section 5.1 that the ability to forecast volatility breaks will produce substantially more accurate volatility forecasts at all forecast horizons by all performance measures used here. In Section 5.2, we show that volatility breaks are widespread and that there is evidence for related and co-breaks in international stock market volatility and among stocks listed in the same stock market. In Section 5.3, we provide some explanations for volatility breaks and in Section 5.4, we outline the difference between volatility breaks and jumps.

5.1 Economic Significance

To demonstrate the impact of breaks in volatility forecast, we estimate and produce forecasts using a B-GJR*(1,1) model assuming that all future volatility break dates and break sizes are known in advance, i.e., we allow the unconditional variance of B-GJR*(1,1) to move in a predictable manner. Figure 3 presents the time series plots of the 1-day forecast alongside the actual realized volatility for GJR(1,1), FIGARCH(1,1), B-GARCH(1,1) and B-GJR*(1,1). From the previous section, GJR and FIGARCH are the best models for short- and long-horizon forecasts, respectively. Figure 4 presents the same information for the 20-day forecast. Apart from the break dates and break sizes used in B-GJR*, all other parameters are estimated from ex ante data only. The GJR, FIGARCH and B-GARCH forecasting results are from Section 4. The GJR and FIGARCH forecasts, presented in Figures 3(a) and 4(a) and Figures 3(b) and 4(b), respectively, are “jagged,” as their volatility persistence parameter values are high, making the forecasts follow very closely the past trend. The two break models, on the other hand, produced volatility forecasts that do not follow the ups and downs of the actual realized volatility, but they mapped exactly to the general shifts in unconditional variance. This is the typical characteristic of a break model that has small volatility persistence parameters. In

\[\text{We have dropped GARCH, VC and RS-GARCH in the analyses here, since we have shown in previous sections that they are inferior in terms of both in-sample fit and out-of-sample forecast.}\]
the case of B-GJR*, the strategy of mapping the volatility mean levels proved to be crucial; it leads to the highest R-square and lowest MAE in both 1- and 20-day forecasts. Unlike the B-GJR* model, the B-GARCH forecasts presented in Figures 3(c) and 4(c) has no foresight information on break dates and break sizes. For this reason, the B-GARCH model suffers from the lag effect in picking up break events, which leads to a high MAE and low R-square as a consequence.

**FIGURE 3 ABOUT HERE**

**FIGURE 4 ABOUT HERE**

Full details of the forecasting results for all other forecast horizons are reported in Table 7. Specifically, Table 7, Panel A evaluates the volatility forecast in levels, whereas Panel B evaluates the volatility forecast in differences to take into account possible non-stationarity in volatility. Panel A shows significant increases in R-square and significant decreases in MAE when volatility break dates and break sizes are known. Detailed results are not reported here, but the B-GJR* model statistically outperformed all other models under both forecast evaluation criterion. The statistics of the “best R-square” models and the “best MAE” models from Table 4 are reproduced in the last two columns of Table 7(A) for ease of comparison. The B-GJR* model with known breaks clearly outperformed these previously best models. In Table 7, Panel B, the R-square of the regression on volatility differences is slightly reduced for the short-horizon forecasts and slightly increased for the long-horizon forecasts, the same as what we noted before in Table 6. Nevertheless, all the R-square and MAE values reported in Panel B for the B-GJR* model again dominate those of the best models reported in Table 6.

**TABLE 7 ABOUT HERE**

In Table 7, Panel A, the $\alpha$ in the regression efficiency test is close to zero and the $\beta$ is close to 1. In Panel B, all the $\beta$ coefficients are less than 1 and all the $\alpha$ coefficients are less than 0, indicating that even the B-GJR* model has a tendency to under-forecast high volatility. This is inevitable in any forecasting models that rely only on historical price information, which forced the models to be backward-looking. The over- and under-forecast phenomenon is best appreciated from the scatterplots presented in Figure 5. The 1-day-ahead forecasting results are presented in the left column of Figure 5, and the 20-day-ahead forecasting results are presented in the right column of the same figure. For the 1-day-ahead forecasts, all models tend to over-forecast in low volatility states, with the dots representing low actual volatility.
lying below the 45° line. On the other hand, all models tend to under-forecast in high volatility states, as shown by the dots representing high actual volatility lying above the 45° line. As the forecast horizon increases, the impact of volatility shocks “averaged out,” and the predictive power improved at both high and low volatility states.

**FIGURE 5 ABOUT HERE**

In reality, it will be difficult to predict volatility breaks themselves, let alone the exact break dates and break sizes. The analyses here serve to highlight the economic significance of volatility breaks and call for more research effort in this direction.

5.2 International Evidence

In this section, we demonstrate that volatility breaks are indeed widespread and could be the potential cause of long memory in volatility. Here, we apply the ICSS break estimation procedure to the 155 financial returns studied in Section 2. To conserve space, we will present here only the results pertaining to the stock markets.

Figure 6 presents the volatility break trends for Japan and three European markets (Germany, France and the UK). Figure 6 shows that some of the volatility breaks are global systemic events, such as the 1987 stock market crash and the 1998 LTCM crisis. On the other hand, the volatility break in July 2002 seems to affect the European countries only. Detailed results not are reported here, but our calculations show that break dependence is stronger between France and Germany and that there is no clear dependence observed between volatility breaks in the UK and those in Japan.

**FIGURE 6 ABOUT HERE**

Figure 7 presents the volatility break trends for three UK stocks and the FTSE small-cap index. All three stocks and the small-cap index appeared to have experienced volatility breaks in the summer of 1994, October 1987 (May in the case of Marks and Spencer) and summer of 1997 to summer of 1998, all of which coincide with well documented worldwide financial crisis. Moreover, the returns and volatility break behavior of the small-cap index suggests that break events are more acute at the aggregate level, as shown by the sharp and more frequent breaks in volatility of the FTSE small-cap index in Figure 7(d). At the index level, idiosyncratic risks are being diversified away, making the undiversifiable systemic risks more pronounced. This does not mean, however, that all break events of all individual series are systemic events.
5.3 Causes of Volatility Breaks

Aggarwal, Inclan and Leal (1999) find that most of the volatility breaks in the emerging markets coincided with country specific economic crises, idiosyncratic political events, as well as worldwide financial crisis. Our own estimation reveals that the duration of volatility breaks in S&P 500 returns ranges from 1 month to 11 years, and the major breaks are all associated with worldwide financial crises.

There have been many studies that relate stock market volatility to macroeconomic conditions (e.g., Officer, 1973; Schwert, 1989; Glosten, Jagannathan and Runkle, 1993; Hamilton and Lin, 1996; David and Veronesi, 2004). But it is Beltratti and Morana (2006) who first establish a formal link between breaks in financial market volatility and volatility breaks in macro variables. Using monthly log variance data for the period January 1970 to September 2000, Beltratti and Morana (2006) document that the break process in S&P 500 volatility is related to volatility breaks in the federal fund rate and M1 growth. After the breaks have been removed, a fractional cointegrating relationship exists linking the “break free” volatility of output growth, the federal fund rate and stock market returns. The “break free” stock market volatility has only small persistent component (15%). Volatility of output growth and, to a lesser extent, the volatility of the federal funds rate and M1 growth affect both the persistent and non-persistent components of S&P 500 volatility.\footnote{Note that in Beltratti and Morana (2006), “structural break” means regime switching only and “long memory” is used generally to refer to volatility persistence.}

In contrast, Kirman and Teyssiere (2001) use a microeconomic model to link herding and swings of opinion with long memory in volatility. The foundation of the model relies on the existence of two groups of agents, chartists and fundamentalists, who may “switch sides” and cause market-wide behavior. The key parameter of the model is $k$, the proportion of fundamentalists, and the probability that agent may change trading strategy independently. The model assumes that an agent will turn into a fundamentalist if he/she observes that $k > 0.5$. Simulations show that if an agent is less likely to behave independently and if $k$ can be observed with little noise, it is sufficient to lead to herding in the extreme and gives rise to bubble-like returns and long memory in absolute and squared returns.
5.4 Distinguishing Breaks and Jumps

A concept that is closely related to breaks in time series is jumps. While the jump diffusion model is not new in finance (see, e.g., Merton, 1976), the argument that jumps in return and jumps in volatility created fat tails and return crashes has received much attention in the finance literature only recently (see, e.g., Chernov, Gallant, Ghysels and Tauchen, 2003; Eraker, Johannes and Polson, 2003; Eraker, 2004; Maheu and McCurdy, 2004). The key difference between volatility jump and volatility break models is their assumptions on the unconditional variance. All the jump studies published to date assume that the unconditional variance is constant.

Moreover, most of these studies find the jump impacts are drastic but short lived. The most common jump-like events that are being investigated are the impact of macro and news releases. Li and Engle (1998) find that scheduled announcements increased volatility in the Treasury futures market, but the impact on volatility has only a small level of persistence. Jones, Lamont and Lumsdaine (1998) find that announcement day shocks in the Treasury bond market produce extra volatility that does not persist at all. Andersen and Bollerslev (1998) again find macroeconomic announcements to produce a large impact on five-minute foreign exchange returns, but the induced effect on volatility is short lived. Flannery and Protoppadakis (2002) fit a GARCH model to macro announcements and stock market returns and find that inflation measures affect the conditional mean, while the real factor measures (e.g., balance of trade, employment and housing starts) enter the conditional variance equation. Along the same line of enquiry, de Goeij and Marquering (2002) use the Asymmetric Dynamic Covariance (ADC) model to show that all the volatility asymmetry in bond markets could be explained by macroeconomic announcements. Macro news do not explain all the volatility asymmetry in the S&P 500 and Nasdaq, however, since the stock markets are also strongly influenced by company news. Separately, Bomfim (2003) also finds unexpected monetary policy produces a short-term and asymmetric impact in stock market volatility.

6 Conclusion

The long-memory characteristic of financial market volatility has been a thorny issue for over 20 years, and we have yet to come to grip on its causes and consequences. There are two schools of thought, although there is no reason to assume that they cannot co-exist. The first school of thought believes that aggregation of volatility components and news arrivals produced the long-memory effect. One volatility dynamic fit for modelling such a process is the fractionally integrated (FI) model that allows the autocorrelation function to decay hyperbolically. The second school of thought believes that financial markets go through periods of structural change producing volatility breaks and regime shifts as a consequence. Volatility is fundamentally short memory between breaks and between regime shifts; the long-memory effect is spurious and is a direct consequence of these breaks and regime shifts. There are fundamental differences between the approaches adopted by the two schools, the most crucial one is the modelling of the unconditional variance. In the case of a break process, the position of the unconditional variance is almost unconstrained in the modelling process.

Whether or not unconditional variance is time varying and, if it is, the way in which it varies have a significant impact on volatility persistence and forecasts. In this paper, we studied three long-memory volatility models—fractionally integrated, break and regime-switching models—all of which are capable of producing a long distributed lags autocorrelation pattern among absolute returns. We studied model fit and volatility forecasting performance using daily returns of S&P 500 over a 32.5 year period from 2 January 1969 to 23 July 2001. In addition, we have studied and estimated the break patterns for a large number of financial time series.

We argue here that volatility break is the major cause of volatility long memory. Many of the volatility break dates and periods are similar across countries and are similar among stocks from the same country. Many of the break episodes coincided with well documented events. In the forecasting exercise, we show that, if one is able to incorporate future volatility breaks into the forecasts, there will be a substantial gain in terms of forecast accuracy and unbiasedness. In practice, it will be difficult to forecast volatility breaks, let alone the exact break dates and break sizes, and current research is limited in informing us how to make such a prediction. In the absence of the ability to predict future volatility breaks, our tests suggest that the fractionally integrated model will be the next best alternative. Hitherto, financial markets behave in a consistent manner in that the probabilities that a volatility break will take place and volatility is non-stationary increase with time. The way in which the FI model
is normally implemented means that the unconditional variance is an exponentially weighted
sum of a long distributed lag of past shocks. This turns out to be a very useful way for coping
with a changing unconditional variance. For short-horizon forecasts of up to 5 days, the short-
memory GJR model is the best. As the forecast horizon extends to beyond 10 to 20 days, it
is important to allow the unconditional variance to change. This poses a significant challenge
to the option pricing literature that up until now has assumed volatility to be stationary with
short memory.

It is now more than two decades since the first publication of the ARCH model by Engle
(1982). But the search for a better dynamic volatility model continues. Engle and Rangel
(2005), for example, propose the spline-GARCH model where the unconditional volatility is
now a deterministic and smoothed function of time. We may view the spline-GARCH as a
smoothed version of the B-GARCH and B-GJR studied here. Spline-GARCH is also similar
to VC, but unlike VC, it allows the unconditional variance to vary through time. As shown
in the empirical study in the previous section, the assumption regarding the behavior of the
unconditional variance has important implications for forecasting. Engle and Rangel (2005)
find that macroeconomic variables such as inflation, output growth and volatility, inflation and
the short rate are related to the unconditional variance of stock returns. It will be interesting for
future research to test if the spline-GARCH model can lead to a better forecasting performance
and if the same set of macroeconomic and uncertainty variables can help to predict volatility
breaks in financial markets.
Appendix A: Quadratic and Bipower Variation

Following Barndorff-Nielsen and Shephard (2004), consider the case where $Y_t$, the log price process, belongs to the Brownian semi-martingale stochastic process with jumps (BSMJ) class:

$$Y_t = \int_0^t a_s ds + \int_0^t \sigma_s dW_s + \sum_{j=1}^{N_t} c_j,$$

(7)

where $ds$ is the drift of the continuous Brownian process $dW_s$ and $c_j$ is the stochastic jump.

Next, define quadratic variation (QV) of $Y$ as

$$[Y]_t = \lim_{n \to \infty} \sum_{j=0}^{n-1} (Y_{t_j+1} - Y_{t_j})^2 = [Y^c]_t + [Y^d]_t,$$

(8)

where $Y^c$ is the continuous part of the local martingale component of $Y$ and $Y^d$ is the discontinuous jump component of $Y$. In practice, a discrete version of (8) is used with log prices observed at an interval of $\delta$. Let $\delta$-returns be given by

$$y_j = Y_{j\delta} - Y_{(j-1)\delta}, \quad j = 1, 2, \ldots, \lfloor t/\delta \rfloor,$$

(9)

with $\lfloor t/\delta \rfloor$ being the integer part of $t/\delta$. If $t$ is in days, then $y_j$ denotes intra-day returns, and the result is the realized quadratic variation:

$$[Y_\delta]_t = \sum_{j=1}^{\lfloor t/\delta \rfloor} y_j^2 = \lim_{\delta \to 0} [Y]_t.$$  

(10)

To separate the two components $[Y^c]_t$ and $[Y^d]_t$, the theory of bipower variation (BPV) from Barndorff-Nielsen and Shephard (2004) is required to derive $[Y^c]_t$. The $[1,1]$-BPV is defined as

$$\{Y\}^{[1,1]}_t = \lim_{\delta \to \infty} \sum_{j=2}^{\lfloor t/\delta \rfloor} |y_{j-1}| |y_j|.$$  

(11)

Again, in practice (11) is estimated from discrete returns:

$$\{Y_\delta\}^{[1,1]}_t = \sum_{j=2}^{\lfloor t/\delta \rfloor} |y_{j-1}| |y_j|.$$  

Barndorff-Nielsen and Shephard (2004) show that if $Y \in$ BSMJ as in (7), provided that $a \equiv 0$ and that the volatility process is independent from the Brownian process, then

$$\{Y\}^{[1,1]}_t = \mu_1^2 \int_0^t \sigma_s^2 ds = \mu_1^2 [Y^c]_t.$$  

29
where \( \mu_1 = E[|u|] \) is the mean of the absolute value of a standard normal random variable, \( u \sim N(0, 1) \). The value of \( \mu_1 \) is \( \sqrt{2/\pi} \approx 0.79788 \). Hence,

\[
[Y^c]_t \doteq \frac{2}{\pi} \{Y\}_t^{[1,1]}.
\]

Equation (10) and (12) can be used to estimate volatility that is due to jumps as follows:

\[
[Y^d]_t = [Y]_t - [Y^c]_t \doteq [Y^\delta]_t - \frac{2}{\pi} \{Y^\delta\}_t^{[1,1]}.
\]

The assumption needed for consistency of (12) rules out volatility leverage effects, which is unfortunately a prominent phenomenon in equity markets. The consistency of (12) is not affected by long memory or breaks in the volatility process.
Acknowledgments

We would like to thank Ole Barndorff-Nielsen, Ralf Becker, Bent Jesper Christensen, Charlotte Christiansen, Mardi Dungey, Chris Orme, Denise Osborn, Enrique Sentana and seminar participants at Economic Studies at Manchester University and Aarhus Business School for helpful comments. Part of this paper was written while Namwon Hyung was a visiting scholar at the Econometric Institute at Erasmus University and Ser-Huang Poon was visiting the Economic Department at University of California in San Diego. They thank the host institutions for their generous support.
References


a flexible coefficient GARCH Model. Manuscript, Pontifical Catholic University of Rio de Janeiro.


Table 1: Sum of autocorrelation coefficients of the first 1,000 lags and the estimation of the differencing parameter for selected financial time series

| Stock Market Indices                                      | No of obs | $\sum \rho(r)$ | $\sum \rho(r^2)$ | $\sum \rho(\ln|r|)$ | $\sum \rho(|Tr|)$ | $d(|r|)$ | $d(r^2)$ | $d(\ln|r|)$ | $d(|Tr|)$ |
|----------------------------------------------------------|-----------|----------------|------------------|---------------------|------------------|---------|---------|----------|---------|
| USA S&P 500 Composite                                    | 9,676     | 35.687         | 3.912            | 27.466              | 40.838           | 0.444*  | 0.179   | 0.486*   | 0.469*  |
| Germany DAX 30 Industrial                                 | 9,634     | 75.571         | 37.102           | 41.890              | 79.186           | 0.468*  | 0.425*  | 0.446*   | 0.466*  |
| Japan NIKKEI 225 Stock Average                           | 8,443     | 89.559         | 23.405           | 84.257              | 95.789           | 0.321   | 0.228   | 0.409*   | 0.318   |
| France CAC 40                                            | 8,276     | 43.310         | 17.467           | 22.432              | 46.539           | 0.404*  | 0.298   | 0.408*   | 0.406*  |
| UK FTSE All Share & FTSE100                              | 8,714     | 30.817         | 12.615           | 18.394              | 33.199           | 0.379   | 0.292   | 0.375*   | 0.395*  |
| Average of 50 stock indices                               |           | 28.345         | 9.743            | 23.502              | 30.288           | 0.407   | 0.294   | 0.406    | 0.415   |

<table>
<thead>
<tr>
<th>Stocks</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Cadbury Schweppes</td>
<td>7,418</td>
<td>48.607</td>
<td>19.236</td>
<td>85.288</td>
<td>50.235</td>
<td>0.462*</td>
<td>0.480*</td>
<td>0.389*</td>
<td>0.457*</td>
</tr>
<tr>
<td>Marks and Spencer Group</td>
<td>7,709</td>
<td>40.635</td>
<td>17.541</td>
<td>67.480</td>
<td>42.575</td>
<td>0.433*</td>
<td>0.399*</td>
<td>0.496*</td>
<td>0.428*</td>
</tr>
<tr>
<td>Shell Transport</td>
<td>8,115</td>
<td>38.947</td>
<td>20.078</td>
<td>44.711</td>
<td>40.035</td>
<td>0.510*</td>
<td>0.501*</td>
<td>0.474*</td>
<td>0.506*</td>
</tr>
<tr>
<td>FTSE Small Cap Index</td>
<td>4,437</td>
<td>25.381</td>
<td>3.712</td>
<td>35.152</td>
<td>28.533</td>
<td>0.331</td>
<td>0.125</td>
<td>0.308</td>
<td>0.352*</td>
</tr>
<tr>
<td>Average of 25 stocks</td>
<td></td>
<td>37.718</td>
<td>14.0213</td>
<td>49.1974</td>
<td>40.0312</td>
<td>0.471</td>
<td>0.373</td>
<td>0.489</td>
<td>0.474</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Exchange Rates</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>US $ to UK £</td>
<td>7,942</td>
<td>56.308</td>
<td>24.652</td>
<td>84.717</td>
<td>57.432</td>
<td>0.488*</td>
<td>0.330</td>
<td>0.549*</td>
<td>0.495*</td>
</tr>
<tr>
<td>Australian $ to UK £</td>
<td>7,859</td>
<td>32.657</td>
<td>0.052</td>
<td>72.572</td>
<td>48.241</td>
<td>0.346</td>
<td>0.072+</td>
<td>0.549*</td>
<td>0.373*</td>
</tr>
<tr>
<td>Mexican Peso to UK £</td>
<td>5,394</td>
<td>9.545</td>
<td>1.501</td>
<td>13.760</td>
<td>14.932</td>
<td>0.105+</td>
<td>0.134+</td>
<td>0.349*</td>
<td>0.249*</td>
</tr>
<tr>
<td>Indonesian Rupiah to UK £</td>
<td>2,964</td>
<td>20.819</td>
<td>4.927</td>
<td>31.509</td>
<td>21.753</td>
<td>0.616*</td>
<td>0.405*</td>
<td>0.629*</td>
<td>0.632*</td>
</tr>
<tr>
<td>Average of 25 exchange rates</td>
<td></td>
<td>14.031</td>
<td>4.078</td>
<td>22.778</td>
<td>16.737</td>
<td>0.362</td>
<td>0.192</td>
<td>0.444</td>
<td>0.394</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Interest Rates</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>US 1 month Eurodollar deposits</td>
<td>8,491</td>
<td>281.799</td>
<td>20.782</td>
<td>327.770</td>
<td>331.877</td>
<td>0.566*</td>
<td>0.105</td>
<td>0.741*</td>
<td>0.670*</td>
</tr>
<tr>
<td>UK Interbank 1-Month</td>
<td>7,448</td>
<td>12.699</td>
<td>0.080</td>
<td>22.901</td>
<td>25.657</td>
<td>0.295</td>
<td>-0.018+</td>
<td>0.346*</td>
<td>0.370*</td>
</tr>
<tr>
<td>Venezuela PAR Brady Bond</td>
<td>3,279</td>
<td>19.236</td>
<td>9.944</td>
<td>32.985</td>
<td>19.800</td>
<td>0.454*</td>
<td>0.301</td>
<td>0.617*</td>
<td>0.475*</td>
</tr>
<tr>
<td>South Korea Overnight Call</td>
<td>2,601</td>
<td>54.693</td>
<td>12.200</td>
<td>57.276</td>
<td>56.648</td>
<td>0.648*</td>
<td>0.496*</td>
<td>0.586*</td>
<td>0.657*</td>
</tr>
<tr>
<td>Average of 34 interest rates</td>
<td></td>
<td>34.987</td>
<td>7.362</td>
<td>44.451</td>
<td>41.306</td>
<td>0.408</td>
<td>0.250</td>
<td>0.493</td>
<td>0.445</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Commodities</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Gold, Bullion, $/troy oz (London fixing) Close</td>
<td>6,536</td>
<td>125.309</td>
<td>39.305</td>
<td>140.747</td>
<td>133.880</td>
<td>0.641*</td>
<td>0.624*</td>
<td>0.561*</td>
<td>0.647*</td>
</tr>
<tr>
<td>Silver Fix (LBM), cash cents/troy oz</td>
<td>7,780</td>
<td>45.504</td>
<td>8.275</td>
<td>88.706</td>
<td>52.154</td>
<td>0.594*</td>
<td>0.249</td>
<td>0.617*</td>
<td>0.666*</td>
</tr>
<tr>
<td>Brent Oil (1 Month Forward) $/barrel</td>
<td>2,389</td>
<td>11.532</td>
<td>5.469</td>
<td>9.882</td>
<td>11.81</td>
<td>0.409*</td>
<td>0.298</td>
<td>0.452*</td>
<td>0.420*</td>
</tr>
<tr>
<td>Average of 21 commodities</td>
<td></td>
<td>19.743</td>
<td>8.073</td>
<td>27.408</td>
<td>20.978</td>
<td>0.361</td>
<td>0.265</td>
<td>0.393</td>
<td>0.367</td>
</tr>
</tbody>
</table>

Note: "Tr" denotes trimmed returns whereby returns in the 0.1% tail take the value of the 0.1% quantile. "d" is the I(d) parameter estimated based on Geweke and Porter-Hudak (1983). * indicates cases where $H_0: d = 0.5$ is not rejected and + indicates cases where $H_0: d = 0$ is not rejected.
Table 2: Sum of 1,000 autocorrelation coefficients for S&P 500 realized volatility and simulated GARCH and GJR processes

<table>
<thead>
<tr>
<th></th>
<th>No of obs</th>
<th>$\sum \rho$ (Quadratic Variation)</th>
<th>$\sum \rho$ (Bipower Variation)</th>
<th>$\sum \rho$ (Jump Component)</th>
</tr>
</thead>
</table>

|                          | No of obs | $\sum \rho(|r|)$ | $\sum \rho(r^2)$ | $\sum \rho(\ln|r|)$ | $\sum \rho(|Tr|)$ |
|--------------------------|-----------|------------------|-------------------|--------------------|-------------------|
| 1,000 simulated realizations of a GARCH process | 10,000    | 1.045            | 1.206             | 0.478              | 1.033             |
| Mean                     |           | (1.099)          | (1.232)           | (0.688)            | (1.086)           |
| Standard Deviation       |           |                  |                   |                    |                   |
| 1,000 simulated realizations of a GJR process | 10,000    | 1.945            | 2.308             | 0.870              | 1.899             |
| Mean                     |           | (1.709)          | (2.048)           | (0.908)            | (1.660)           |
| Standard Deviation       |           |                  |                   |                    |                   |

Note: "Tr" denotes trimmed returns whereby returns in the 0.01% tail take the value of the 0.01% quantile.

The simulated GARCH process is

$$\varepsilon_t = z_t \sqrt{h_t} \quad \varepsilon_t \sim N(0, 1)$$

$$h_t = (1 - 0.96 - 0.02) + 0.96 h_{t-1} + 0.02 \varepsilon_{t-1}^2$$

The simulated GJR process is

$$\varepsilon_t = z_t \sqrt{h_t} \quad \varepsilon_t \sim N(0, 1)$$

$$h_t = (1 - 0.9 - 0.03 - 0.5 \times 0.09) + 0.9 h_{t-1} + 0.03 \varepsilon_{t-1}^2 + 0.09 D_{t-1} \varepsilon_{t-1}^2$$

$$D_t = \begin{cases} 1 & \text{for } \varepsilon_t < 0 \\ 0 & \text{otherwise} \end{cases}$$
Table 3: Models for daily returns on S&P 500 index from 2 January 1969 to 18 January 1991

\[ r_t = \mu + \varepsilon_t, \quad \varepsilon_t = z_t \sqrt{h_t}, \quad z_t \sim N(0,1). \]

GARCH(1,1) \[ h_t = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \]

GJR-GARCH(1,1) \[ h_t = \omega + \alpha_1 \varepsilon_{t-1}^2 I_{[\varepsilon_{t-1} \geq 0]} + \gamma_1 \varepsilon_{t-1}^2 I_{[\varepsilon_{t-1} < 0]} + \beta_1 h_{t-1} \]

FIGARCH(1,1) \[ h_t = \omega + \left[ 1 - \beta_1 \right] L - (1 - \phi_1 L)(1 - L)^d \] \[ \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \]

B-GARCH(1,1) \[ h_t = \omega_0 D_0 + \cdots + \omega_R D_R + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \]

VC \[ h_t - \omega_t = \alpha_1 (\varepsilon_{t-1}^2 - \omega_{t-1}) + \beta_1 (h_{t-1} - \omega_{t-1}), \omega_{t} = \omega + \rho \omega_{t-1} + \varphi (\varepsilon_{t-1}^2 - h_{t-1}) \]

RS-GARCH \[ h_t^R = \alpha_{S_{t-1}} + \alpha_{1S_{t-1}} \varepsilon_{t-1}^2 + \beta_{1S_{t-1}} h_{t-1} \]

where \( r_t \) is S&P daily returns at time \( t \), \( D_0, \ldots, D_R \) are dummy variables taking the value of 1 for a particular regime and 0 otherwise.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>GARCH</th>
<th>GJR</th>
<th>FIGARCH</th>
<th>B-GARCH</th>
<th>VC</th>
<th>RS-GARCH</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mu )</td>
<td>0.037*</td>
<td>0.024*</td>
<td>0.037*</td>
<td>0.028</td>
<td>0.037*</td>
<td>0.037*</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.012)</td>
<td>(0.011)</td>
<td>(0.017)</td>
<td>(0.011)</td>
<td>(0.010)</td>
</tr>
<tr>
<td>( \omega )</td>
<td>0.017*</td>
<td>0.018*</td>
<td>0.016*</td>
<td>0.011*</td>
<td>0.330*</td>
<td>0.330*</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.002)</td>
<td>(0.048)</td>
<td>(0.048)</td>
</tr>
<tr>
<td>( \alpha_1 )</td>
<td>0.078*</td>
<td>0.108*</td>
<td>0.086*</td>
<td>0.095*</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.010)</td>
<td>(0.017)</td>
<td>(0.017)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>( \gamma_1 )</td>
<td>0.035*</td>
<td>0.370*</td>
<td>0.040</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.064)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \phi_1 )</td>
<td></td>
<td>0.545*</td>
<td>0.064</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>0.905*</td>
<td>0.910*</td>
<td>0.769*</td>
<td>0.036</td>
<td>0.055*</td>
<td>0.276*</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.009)</td>
<td>(0.035)</td>
<td>(0.264)</td>
<td>(0.006)</td>
<td>(0.088)</td>
</tr>
<tr>
<td>( \omega_2 )</td>
<td></td>
<td>0.818*</td>
<td>(0.156)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \alpha_2 )</td>
<td></td>
<td>0.158*</td>
<td>(0.027)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \beta_2 )</td>
<td></td>
<td>0.537*</td>
<td>(0.079)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Log-likelihood | -7127.25 | -7095.84 | -7108.30 | -6964.10 | -7111.68 | -7068.12
Volatility persistent | 0.983 | 0.982 | 0.122 | 0.991 | 0.276 (L) | 0.695 (H)
Unconditional variance | 1.000 | 0.9730 | 1.000 | 0.8180 |

Notes: Parameters are estimated based on quasi maximum likelihood function. Bollerslev-Wooldridge (1992) robust standard errors are shown in parentheses. For the regime-switching model, the probability of staying in the low volatility regime is \( P_{11} = 0.991 \) (0.002) and the probability of staying in the high volatility regime is \( P_{22} = 0.978 \) (0.006) with standard errors shown in parentheses. The number of breaks in the B-GARCH model is \( R = 27 \). The volatility persistence measure is calculated as \( \alpha_1 + \beta_1 \) for GARCH, B-GARCH and each regime in RS-GARCH, \( \alpha_1 + 0.5 \gamma_1 + \beta_1 \) for GJR, and \( (\alpha_1 + \beta_1)(1 - \beta) \) for VC.
Table 4: Forecast evaluation for S&P 500 return volatility from 21 January 1991 to 23 July 2001 based on realized volatility calculated using intraday returns

\[
MAE = \frac{1}{m} \sum_{i=0}^{m-1} \frac{1}{L} \sum_{s=1}^{L} \left( \sqrt{\sum_{t=1}^{s} h_{t+s}^2} - \sqrt{\sum_{t=1}^{s} \hat{h}_{t+s}^2} \right) 
= \alpha + \beta \sqrt{\sum_{t=1}^{s} \hat{h}_{t+s}^2} + \upsilon_t.
\]

<table>
<thead>
<tr>
<th>Prediction horizon, s</th>
<th>Joint test</th>
<th>MAE</th>
<th>R-square</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GARCH(1,1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>307.371</td>
<td>0.000</td>
<td>-0.061</td>
</tr>
<tr>
<td>5</td>
<td>41.615</td>
<td>0.000</td>
<td>-0.034</td>
</tr>
<tr>
<td>10</td>
<td>20.587</td>
<td>0.000</td>
<td>0.033</td>
</tr>
<tr>
<td>20</td>
<td>10.950</td>
<td>0.004</td>
<td>0.162</td>
</tr>
<tr>
<td>60</td>
<td>6.989</td>
<td>0.157</td>
<td>0.520</td>
</tr>
<tr>
<td>120</td>
<td>1.087</td>
<td>0.581</td>
<td>0.064</td>
</tr>
<tr>
<td>GJR(1,1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>364.135</td>
<td>0.000</td>
<td>-0.073</td>
</tr>
<tr>
<td>5</td>
<td>58.894</td>
<td>0.000</td>
<td>-0.028</td>
</tr>
<tr>
<td>10</td>
<td>31.306</td>
<td>0.000</td>
<td>0.016</td>
</tr>
<tr>
<td>20</td>
<td>17.264</td>
<td>0.000</td>
<td>0.084</td>
</tr>
<tr>
<td>60</td>
<td>5.380</td>
<td>0.068</td>
<td>-0.124</td>
</tr>
<tr>
<td>120</td>
<td>2.391</td>
<td>0.303</td>
<td>-1.982</td>
</tr>
<tr>
<td>FIGARCH(1,1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>323.125</td>
<td>0.000</td>
<td>-0.032</td>
</tr>
<tr>
<td>5</td>
<td>49.551</td>
<td>0.000</td>
<td>0.016</td>
</tr>
<tr>
<td>10</td>
<td>25.989</td>
<td>0.000</td>
<td>0.083</td>
</tr>
<tr>
<td>20</td>
<td>13.459</td>
<td>0.000</td>
<td>0.158</td>
</tr>
<tr>
<td>60</td>
<td>4.507</td>
<td>0.010</td>
<td>0.273</td>
</tr>
<tr>
<td>120</td>
<td>1.997</td>
<td>0.368</td>
<td>-0.050</td>
</tr>
<tr>
<td>B-GARCH(1,1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>544.439</td>
<td>0.000</td>
<td>0.121</td>
</tr>
<tr>
<td>5</td>
<td>94.045</td>
<td>0.000</td>
<td>0.395</td>
</tr>
<tr>
<td>10</td>
<td>51.919</td>
<td>0.000</td>
<td>0.643</td>
</tr>
<tr>
<td>20</td>
<td>31.889</td>
<td>0.000</td>
<td>1.077</td>
</tr>
<tr>
<td>60</td>
<td>21.901</td>
<td>0.000</td>
<td>2.518</td>
</tr>
<tr>
<td>120</td>
<td>18.345</td>
<td>0.000</td>
<td>3.807</td>
</tr>
<tr>
<td>VC(1,1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>393.648</td>
<td>0.000</td>
<td>-0.059</td>
</tr>
<tr>
<td>5</td>
<td>61.870</td>
<td>0.000</td>
<td>-0.285</td>
</tr>
<tr>
<td>10</td>
<td>31.978</td>
<td>0.000</td>
<td>0.022</td>
</tr>
<tr>
<td>20</td>
<td>18.494</td>
<td>0.000</td>
<td>0.146</td>
</tr>
<tr>
<td>60</td>
<td>8.019</td>
<td>0.018</td>
<td>0.502</td>
</tr>
<tr>
<td>120</td>
<td>4.012</td>
<td>0.135</td>
<td>0.028</td>
</tr>
<tr>
<td>RS-GARCH(1,1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>390.709</td>
<td>0.000</td>
<td>-0.103</td>
</tr>
<tr>
<td>5</td>
<td>47.799</td>
<td>0.000</td>
<td>-0.218</td>
</tr>
<tr>
<td>10</td>
<td>18.487</td>
<td>0.000</td>
<td>-0.395</td>
</tr>
<tr>
<td>20</td>
<td>8.837</td>
<td>0.012</td>
<td>-0.919</td>
</tr>
<tr>
<td>60</td>
<td>6.957</td>
<td>0.031</td>
<td>-4.543</td>
</tr>
<tr>
<td>120</td>
<td>5.613</td>
<td>0.060</td>
<td>-13.242</td>
</tr>
<tr>
<td>ES</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>365.846</td>
<td>0.000</td>
<td>0.079</td>
</tr>
<tr>
<td>5</td>
<td>61.051</td>
<td>0.000</td>
<td>0.285</td>
</tr>
<tr>
<td>10</td>
<td>35.176</td>
<td>0.000</td>
<td>0.500</td>
</tr>
<tr>
<td>20</td>
<td>15.890</td>
<td>0.000</td>
<td>2.145</td>
</tr>
<tr>
<td>60</td>
<td>13.561</td>
<td>0.001</td>
<td>3.444</td>
</tr>
<tr>
<td>120</td>
<td>6.436</td>
<td>0.040</td>
<td>2.701</td>
</tr>
<tr>
<td>RW</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>3654.705</td>
<td>0.000</td>
<td>0.585</td>
</tr>
<tr>
<td>5</td>
<td>333.364</td>
<td>0.000</td>
<td>0.849</td>
</tr>
<tr>
<td>10</td>
<td>126.437</td>
<td>0.000</td>
<td>1.036</td>
</tr>
<tr>
<td>20</td>
<td>52.117</td>
<td>0.000</td>
<td>1.326</td>
</tr>
<tr>
<td>60</td>
<td>15.369</td>
<td>0.000</td>
<td>2.401</td>
</tr>
<tr>
<td>120</td>
<td>6.436</td>
<td>0.040</td>
<td>2.701</td>
</tr>
</tbody>
</table>

Notes: BGARCH is GARCH(1,1) with occasional breaks in variance, VC is the Engle and Lee (1999) volatility component model, RS-GARCH is the regime switching model, ES is exponential smoothing and RW is the random walk model. Each of the daily forecasts was made using parameters estimated using returns on previous 5,550 trading days. The best model for each forecast horizon is in <>. * denotes cases where the null hypothesis of equal forecast accuracy against the <best model> is rejected at the 5% significance level using the Diebold and Mariano test.
Table 5: Forecast evaluation for S&P 500 return volatility from 21 January 1991 to 23 July 2001 based on volatility proxy calculated as daily squared returns

<table>
<thead>
<tr>
<th>Prediction horizon, s</th>
<th>GARCH(1,1)</th>
<th>GJR(1,1)</th>
<th>FIGARCH(1,1)</th>
<th>B-GARCH(1,1)</th>
<th>VC(1,1)</th>
<th>RS-GARCH(1,1)</th>
<th>ES</th>
<th>RW</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MAE 346.79</td>
<td>369.58</td>
<td>359.93</td>
<td>493.60</td>
<td>397.78</td>
<td>393.22</td>
<td>376.94</td>
<td>1854.44</td>
</tr>
<tr>
<td>5</td>
<td>17.66</td>
<td>19.74</td>
<td>21.25</td>
<td>52.79</td>
<td>28.31</td>
<td>21.09</td>
<td>28.42</td>
<td>223.80</td>
</tr>
<tr>
<td>10</td>
<td>5.67</td>
<td>4.73</td>
<td>7.60</td>
<td>16.46</td>
<td>11.29</td>
<td>6.15</td>
<td>6.30</td>
<td>10.06</td>
</tr>
<tr>
<td>20</td>
<td>0.25</td>
<td>0.58</td>
<td>0.44</td>
<td>11.47</td>
<td>1.85</td>
<td>7.67</td>
<td>8.30</td>
<td>3.45</td>
</tr>
<tr>
<td>60</td>
<td>0.05</td>
<td>0.25</td>
<td>0.30</td>
<td>0.98</td>
<td>0.98</td>
<td>8.22</td>
<td>1.06</td>
<td>8.22</td>
</tr>
<tr>
<td>120</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: BGARCH is GARCH(1,1) with occasional breaks in variance, VC is the Engle and Lee (1999) volatility component model, RS-GARCH is the regime switching model, ES is exponential smoothing and RW is the random walk model. Each of the daily forecasts was made using parameters estimated using returns on previous 5,550 trading days. The best model for each forecast horizon is in <>. * denotes cases where the null hypothesis of equal forecast accuracy against the <best model> is rejected at the 5% significance level using the Diebold and Mariano test.
<table>
<thead>
<tr>
<th>Prediction horizon, s</th>
<th>Joint-test (Std. err.)</th>
<th>α (t-ratio)</th>
<th>β (t-ratio)</th>
<th>MAE</th>
<th>R-square</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>GARCH(1,1)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>636.433 0.000</td>
<td>-0.081</td>
<td>-11.366</td>
<td>0.663</td>
<td>-17.462</td>
</tr>
<tr>
<td>5</td>
<td>228.840 0.000</td>
<td>-0.158</td>
<td>-3.042</td>
<td>0.745</td>
<td>-10.266</td>
</tr>
<tr>
<td>10</td>
<td>131.308 0.000</td>
<td>-0.216</td>
<td>-2.153</td>
<td>0.785</td>
<td>-7.616</td>
</tr>
<tr>
<td>20</td>
<td>69.837 0.000</td>
<td>-0.345</td>
<td>-1.313</td>
<td>0.861</td>
<td>-4.015</td>
</tr>
<tr>
<td>60</td>
<td>19.643 0.009</td>
<td>-0.360</td>
<td>-2.107</td>
<td>0.871</td>
<td>-2.863</td>
</tr>
<tr>
<td>120</td>
<td>9.498 0.009</td>
<td>-0.360</td>
<td>-0.735</td>
<td>0.871</td>
<td>-2.635</td>
</tr>
<tr>
<td><strong>GJR(1,1)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>492.069 0.000</td>
<td>-0.097</td>
<td>-13.791</td>
<td>0.772</td>
<td>-11.032</td>
</tr>
<tr>
<td>5</td>
<td>158.472 0.000</td>
<td>-0.156</td>
<td>-4.557</td>
<td>0.809</td>
<td>-8.060</td>
</tr>
<tr>
<td>10</td>
<td>100.056 0.000</td>
<td>-0.205</td>
<td>-3.042</td>
<td>0.821</td>
<td>-6.200</td>
</tr>
<tr>
<td>20</td>
<td>55.060 0.000</td>
<td>-0.297</td>
<td>-2.153</td>
<td>0.840</td>
<td>-5.996</td>
</tr>
<tr>
<td>60</td>
<td>15.253 0.000</td>
<td>-0.549</td>
<td>-1.313</td>
<td>0.861</td>
<td>-4.015</td>
</tr>
<tr>
<td>120</td>
<td>8.986 0.011</td>
<td>-0.360</td>
<td>-0.735</td>
<td>0.871</td>
<td>-2.863</td>
</tr>
<tr>
<td><strong>FIGARCH(1,1)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>475.466 0.000</td>
<td>-0.087</td>
<td>-12.265</td>
<td>0.737</td>
<td>-12.645</td>
</tr>
<tr>
<td>5</td>
<td>184.343 0.000</td>
<td>-0.133</td>
<td>-5.015</td>
<td>0.758</td>
<td>-11.151</td>
</tr>
<tr>
<td>10</td>
<td>109.719 0.000</td>
<td>-0.175</td>
<td>-3.482</td>
<td>0.783</td>
<td>-9.019</td>
</tr>
<tr>
<td>20</td>
<td>56.442 0.000</td>
<td>-0.247</td>
<td>-2.567</td>
<td>0.821</td>
<td>-5.996</td>
</tr>
<tr>
<td>60</td>
<td>12.854 0.002</td>
<td>-0.549</td>
<td>-1.957</td>
<td>0.873</td>
<td>-2.961</td>
</tr>
<tr>
<td>120</td>
<td>6.663 0.036</td>
<td>-0.659</td>
<td>-1.096</td>
<td>0.832</td>
<td>-2.509</td>
</tr>
<tr>
<td><strong>B-GARCH(1,1)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>994.930 0.000</td>
<td>-0.076</td>
<td>-10.258</td>
<td>0.552</td>
<td>-24.550</td>
</tr>
<tr>
<td>5</td>
<td>452.383 0.000</td>
<td>-0.109</td>
<td>-3.984</td>
<td>0.580</td>
<td>-19.175</td>
</tr>
<tr>
<td>10</td>
<td>254.676 0.000</td>
<td>-0.137</td>
<td>-2.527</td>
<td>0.604</td>
<td>-14.783</td>
</tr>
<tr>
<td>20</td>
<td>141.207 0.000</td>
<td>-0.174</td>
<td>-1.608</td>
<td>0.635</td>
<td>-7.540</td>
</tr>
<tr>
<td>60</td>
<td>65.029 0.000</td>
<td>-0.229</td>
<td>-0.940</td>
<td>0.698</td>
<td>-7.722</td>
</tr>
<tr>
<td>120</td>
<td>46.078 0.000</td>
<td>-0.221</td>
<td>-0.806</td>
<td>0.743</td>
<td>-6.350</td>
</tr>
<tr>
<td><strong>VC(1,1)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>572.129 0.000</td>
<td>-0.097</td>
<td>-13.439</td>
<td>0.728</td>
<td>-13.116</td>
</tr>
<tr>
<td>5</td>
<td>225.581 0.000</td>
<td>-0.157</td>
<td>-5.783</td>
<td>0.736</td>
<td>-11.893</td>
</tr>
<tr>
<td>10</td>
<td>133.544 0.000</td>
<td>-0.210</td>
<td>-3.042</td>
<td>0.757</td>
<td>-9.685</td>
</tr>
<tr>
<td>20</td>
<td>73.586 0.000</td>
<td>-0.301</td>
<td>-1.559</td>
<td>0.793</td>
<td>-7.311</td>
</tr>
<tr>
<td>60</td>
<td>23.803 0.002</td>
<td>-0.582</td>
<td>-2.227</td>
<td>0.867</td>
<td>-3.896</td>
</tr>
<tr>
<td>120</td>
<td>12.686 0.002</td>
<td>-0.825</td>
<td>-1.723</td>
<td>0.881</td>
<td>-2.746</td>
</tr>
<tr>
<td><strong>RS-GARCH(1,1)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>578.851 0.000</td>
<td>-0.099</td>
<td>-13.517</td>
<td>0.726</td>
<td>-12.918</td>
</tr>
<tr>
<td>5</td>
<td>208.469 0.000</td>
<td>-0.136</td>
<td>-5.194</td>
<td>0.736</td>
<td>-11.768</td>
</tr>
<tr>
<td>10</td>
<td>124.961 0.000</td>
<td>-0.194</td>
<td>-2.952</td>
<td>0.741</td>
<td>-9.971</td>
</tr>
<tr>
<td>20</td>
<td>65.052 0.000</td>
<td>-0.301</td>
<td>-1.559</td>
<td>0.757</td>
<td>-7.540</td>
</tr>
<tr>
<td>60</td>
<td>19.232 0.000</td>
<td>-0.092</td>
<td>-0.290</td>
<td>0.769</td>
<td>-4.304</td>
</tr>
<tr>
<td>120</td>
<td>12.384 0.002</td>
<td>0.057</td>
<td>0.088</td>
<td>0.743</td>
<td>-3.504</td>
</tr>
<tr>
<td><strong>ES</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>690.356 0.000</td>
<td>-0.069</td>
<td>-9.700</td>
<td>0.616</td>
<td>-20.421</td>
</tr>
<tr>
<td>5</td>
<td>267.392 0.000</td>
<td>-0.094</td>
<td>-3.450</td>
<td>0.669</td>
<td>-14.918</td>
</tr>
<tr>
<td>10</td>
<td>151.397 0.000</td>
<td>-0.112</td>
<td>-2.120</td>
<td>0.696</td>
<td>-11.544</td>
</tr>
<tr>
<td>20</td>
<td>80.414 0.000</td>
<td>-0.138</td>
<td>-1.341</td>
<td>0.734</td>
<td>-8.551</td>
</tr>
<tr>
<td>60</td>
<td>26.315 0.000</td>
<td>-0.159</td>
<td>-0.633</td>
<td>0.819</td>
<td>-4.969</td>
</tr>
<tr>
<td>120</td>
<td>13.387 0.001</td>
<td>-0.056</td>
<td>-0.197</td>
<td>0.856</td>
<td>-3.559</td>
</tr>
<tr>
<td><strong>RW</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>4045.700 0.000</td>
<td>0.012</td>
<td>1.467</td>
<td>0.114</td>
<td>62.318</td>
</tr>
<tr>
<td>5</td>
<td>702.305 0.000</td>
<td>0.025</td>
<td>1.096</td>
<td>0.444</td>
<td>26.445</td>
</tr>
<tr>
<td>10</td>
<td>307.932 0.000</td>
<td>0.001</td>
<td>0.027</td>
<td>0.567</td>
<td>-17.388</td>
</tr>
<tr>
<td>20</td>
<td>139.719 0.000</td>
<td>0.049</td>
<td>0.517</td>
<td>0.661</td>
<td>-11.637</td>
</tr>
<tr>
<td>60</td>
<td>38.328 0.000</td>
<td>-0.074</td>
<td>0.277</td>
<td>0.743</td>
<td>-6.099</td>
</tr>
<tr>
<td>120</td>
<td>16.176 0.000</td>
<td>-0.042</td>
<td>-0.107</td>
<td>0.815</td>
<td>-3.960</td>
</tr>
</tbody>
</table>

Notes: BGARCH is GARCH(1,1) with occasional breaks in variance, VC is the Engle and Lee (1999) volatility component model, RS-GARCH is the regime switching model, ES is exponential smoothing and RW is the random walk model. Each of the daily forecasts was made using parameters estimated using returns on previous 5,550 trading days. The best model for each forecast horizon is in <>. * denotes cases where the null hypothesis of equal forecast accuracy against the <best model> is rejected at the 5% significance level using the Diebold and Mariano test.
Table 7: Forecast evaluation for S&P 500 return volatility from 21 January 1991 to 23 July 2001 based on B-GJR(1,1) model assuming breaks dates in the forecasting period are known and with realized volatility calculated using intraday return

<table>
<thead>
<tr>
<th>Prediction horizon</th>
<th>Joint-test (Std. err.)</th>
<th>α (t-ratio)</th>
<th>β (t-ratio)</th>
<th>R-square</th>
<th>MAE</th>
<th>Best R-square models in Table 4 &amp; R-square values</th>
<th>Best RAE models in Table 4 &amp; MAE values</th>
</tr>
</thead>
<tbody>
<tr>
<td>B-GJR(1,1)</td>
<td>1</td>
<td>312.39</td>
<td>-0.016</td>
<td>0.968</td>
<td>0.899</td>
<td>-5.943</td>
<td>0.515</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>48.47</td>
<td>0.024</td>
<td>0.337</td>
<td>0.905</td>
<td>-2.937</td>
<td>0.637</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>23.53</td>
<td>0.053</td>
<td>0.413</td>
<td>0.907</td>
<td>-2.184</td>
<td>0.681</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>13.45</td>
<td>0.057</td>
<td>0.266</td>
<td>0.921</td>
<td>-1.598</td>
<td>0.747</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>6.86</td>
<td>0.032</td>
<td>0.150</td>
<td>0.939</td>
<td>-1.087</td>
<td>0.864</td>
</tr>
<tr>
<td></td>
<td>120</td>
<td>5.52</td>
<td>0.063</td>
<td>0.233</td>
<td>0.361</td>
<td>-1.157</td>
<td>0.916</td>
</tr>
</tbody>
</table>

A. \[
\sqrt{\sum_{i=1}^{s} h_{t+i}} = \alpha + \beta \sqrt{\sum_{j=1}^{s} \hat{h}_{t+j}} + \nu_t .
\]

B-GJR(1,1)

| B-GJR(1,1) | 1 | 353.14 | 0.000 | -0.088 | -13.156 | 0.829 | -8.501 | 0.391 | 0.252 | 0.257 | 0.301 | GJR | 0.464 | GJR | 0.268 |
| B-GJR(1,1) | 5 | 109.00 | 0.000 | -0.130 | -5.341 | 0.847 | -7.731 | 0.572 | 0.425 | 0.381 | 0.556 | GJR | 0.518 | GJR | 0.476 |
| B-GJR(1,1) | 10 | 59.69 | 0.000 | -0.163 | -3.651 | 0.869 | -6.106 | 0.653 | 0.544 | 0.423 | 0.752 | FIGARCH | 0.520 | FIGARCH | 0.645 |
| B-GJR(1,1) | 20 | 25.75 | 0.000 | -0.219 | -2.900 | 0.919 | -3.704 | 0.755 | 0.672 | 0.452 | 1.051 | FIGARCH | 0.510 | FIGARCH | 0.890 |
| B-GJR(1,1) | 60 | 6.13 | 0.047 | -0.346 | -2.231 | 0.983 | -0.838 | 0.899 | 0.815 | 0.483 | 1.789 | FIGARCH | 0.468 | ES | 1.594 |
| B-GJR(1,1) | 120 | 4.15 | 0.125 | -0.446 | -1.840 | 0.988 | -0.723 | 0.944 | 0.947 | 0.529 | 2.395 | FIGARCH | 0.498 | RW | 2.119 |

B. \[
\left( \sqrt{\sum_{i=1}^{s} h_{t+i}} - \sqrt{sh_t} \right) = \alpha + \beta \left( \sqrt{\sum_{j=1}^{s} \hat{h}_{t+j}} - \sqrt{sh_t} \right) + \nu_t .
\]

Best R-square models in Table 6 & R-square values

| B-GJR(1,1) | 1 | 353.14 | 0.000 | -0.088 | -13.156 | 0.829 | -8.501 | 0.391 | 0.252 | 0.257 | 0.301 | GJR | 0.345 |
| B-GJR(1,1) | 5 | 109.00 | 0.000 | -0.130 | -5.341 | 0.847 | -7.731 | 0.572 | 0.425 | 0.381 | 0.556 | GJR | 0.462 |
| B-GJR(1,1) | 10 | 59.69 | 0.000 | -0.163 | -3.651 | 0.869 | -6.106 | 0.653 | 0.544 | 0.423 | 0.752 | FIGARCH | 0.495 |
| B-GJR(1,1) | 20 | 25.75 | 0.000 | -0.219 | -2.900 | 0.919 | -3.704 | 0.755 | 0.672 | 0.452 | 1.051 | FIGARCH | 0.539 |
| B-GJR(1,1) | 60 | 6.13 | 0.047 | -0.346 | -2.231 | 0.983 | -0.838 | 0.899 | 0.815 | 0.483 | 1.789 | FIGARCH | 0.618 |
| B-GJR(1,1) | 120 | 4.15 | 0.125 | -0.446 | -1.840 | 0.988 | -0.723 | 0.944 | 0.947 | 0.529 | 2.395 | FIGARCH | 0.660 |

Notes: Each of the daily forecasts was made using parameters estimated using returns on previous 5,550 trading days. The actual volatility is calculated after the breaks in volatility have been removed, assuming that the break dates and break sizes are known. With this assumption and adjustment to actual volatility calculation, the B-GJR model outperformed all other volatility models at all forecasting horizons.
Figure 2: Autocorrelation coefficients of first 100 lags of absolute returns and the first 120 days volatility point forecasts using returns simulated from an N(0,1) distribution and the parameters from six volatility models estimated and reported in Table 3.
Figure 3: 1-day-ahead volatility forecasts against actual intraday realized volatility

(a) GJR(1,1)  
MAE = 0.268, R-square = 0.464

(b) FIGARCH(1,1)  
MAE = 0.269, R-square = 0.440

(c) B-GARCH(1,1)  
MAE = 0.301, R-square = 0.334

(d) B-GJR*(1,1)  
MAE = 0.252, R-square = 0.515

Note: Forecasts from B-GJR* assume break dates and break sizes are known.
Note: Forecasts from B-GJR* assume break dates and break sizes are known.
Figure 5: Scatter plot of 1- and 20-day actual intraday realized volatility against model forecasts

Notes: Forecasts from B-GJR* assume that the break dates and break sizes are known. The numbers in brackets are estimates of the intercept and slope coefficients for (5). The solid line indicates a 45-degrees line through the data points.
Figure 6: Stock market returns and volatility breaks estimated using the ICSS procedure

(a) Germany Dax 30 Industrial

(b) Japan Nikkei 225 Stock Average

(c) France CAC40

(d) UK FTSE All Share & FTSE 100
Figure 7: Volatility breaks among UK stock returns and small stock index using the ICSS procedure