Modeling Asset Volatility

Jonathan Kinlay

2002

`

Contents

Introduction

Volatility is a central concept in finance, with applications from asset pricing to risk management. Early asset pricing models (e.g. Black and Scholes, 1973) were based on the assumption of constant volatility, but it is now widely accepted that volatility is time varying. A host of econometric models have been developed to represent this and other stylized facts about volatility processes and applied extensively in theoretical and empirical finance. Thus we have the ARCH, GARCH and EGARCH-family of models representing conditional heteroscedasticity (e.g. Engle, 1982; Bollerslev, 1987; Nelson, 1990), as well as a plethora of stochastic volatility models (e.g. Hull and White, 1987; Heston, 1993). All suffer from drawbacks. In the case of GARCH models, while the phenomenon of volatility clustering is typically well represented, other features which have been the focus of more recent empirical research are less adequately described. These include the characteristic of long-memory, typified by a pattern of slowly-decaying autocorrelations, as well as the phenomenon of volatility asymmetry, in which positive and negative shocks induce differing degrees of asset volatility. We shall examine some current attempts to capture this and other important features of asset volatility processes within the GARCH framework.

For stochastic volatility models a rather different problem predominates, that of estimation. Thus the Gaussian quasi-maximum likelihood estimation (QMLE) of Ruiz (1994) and Harvey, Ruiz and Shephard (1994) suffers from the problem that stochastic volatility models are highly non-Gaussian. Andersen and Sorensen (1997) point to the root of the difficulty: standard volatility metrics such as log absolute or squared returns are contaminated by a highly non-Gaussian source of error which produces very inefficient Gaussian GLME estimators. Similarly inefficient are the estimators produced by variants of the generalized method of moments (GMM) whether obtained though simulation (e.g. Duffie and Singleton, 1993) or analytically (Singleton 1997). We follow more recent attempts to find volatility metrics which address the efficiency problem while preserving the appeal of stochastic volatility models.

Multivariate extensions to univariate models enable researchers to address the question of whether an increase in volatility in one market induces additional volatility in another. The concept of cointegration was introduced by Granger 1986) and Engle and Granger (1987) and extended by Baillie and Bollerslev (1994) to cover fractionally integrated processes which exhibit the property of long memory. We consider applications of these more recent concepts to empirical research into volatility co-movement.

Volatility Metrics and their Properties

Standard Volatility Proxies

ABD consider a volatility proxy that is a statistic $f(s_{iH,(i+1)H})$ of the continuous sample path $s_{iH,(i+1)H}$ of the log asset price between times iH and (i+1)H. If the statistic is homogeneous in some power γ of volatility, then it can be written as:

$$f(s_{iH,(i+1)H}) = \sigma_{iH}^{\gamma} f(s^{*}_{iH,(i+1)H}) \, ,$$

which implies that:

$$\ln\left|f(s_{iH,(i+1)H})\right| = \gamma \ln \sigma_{iH} + \ln\left|f(s^{*}_{iH,(i+1)H})\right|$$

where s* is the continuous sample path of the standardized diffusion with unit volatility.

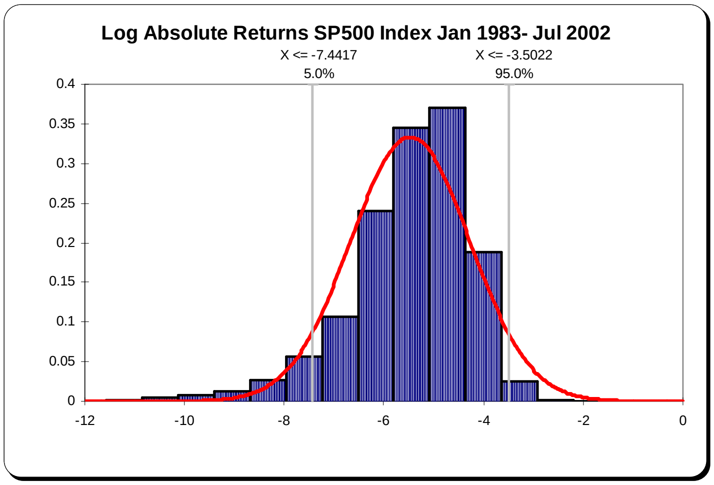
This makes clear that the statistic f(•) is a noisy volatility metric – the first term is proportional to log volatility while the second term is a measurement error. The more variable the measurement error, the less precise are inferences regarding the log volatility and its dynamics.

Theoretical models tend to use either absolute or squared returns as volatility proxies, which in the notation above gives:

$$\ln\left|f(s_{iH,(i+1)H})\right| = \gamma \ln \sigma_{iH} + \gamma \ln\left|s^{*}_{(i+1)H} - s^{*}_{iH}\right|$$

where γ = 1 or 2, depending on whether we are using absolute or squared returns. As ABD point out, γ only scales the volatility proxy and does not affect the distribution. Consequently they focus, without loss of generality, on absolute returns. Using a result from Karatzas and Shreve (1991) ABD are able to demonstrate the non-Gaussian properties on log absolute returns, for which the skewness and kurtosis are –1.53 and 6.93. The figure below shows the empirical distribution of log absolute returns on the SP500 index from January 1983 to July 2002. The distribution is severely skewed (-1.08) and has much higher kurtosis

(5.0) the standard Normal distribution. Jacquier, Polson and Rossi (1994), Andersen and Sorensen (1997) and Kim, Sheppard and Chib (1998) argue that, as a result, Gaussian QMLE with these traditional volatility proxies is highly inefficient and often severely biased in finite samples.



SP500 Index: Log Absolute Returns

January 1983 – June 2002

Realized Volatility

Andersen, Bollerslev, Diebold and Ebens, 2000 (hereafter ABDE) introduce the idea of using high frequency returns to construct estimates of the ex-post realized volatility. For intuition the authors point to the work of Merton (1980) and Nelson (1992) for a continuous time diffusion process, which indicates that the diffusion coefficients can be determined arbitrarily well with sufficiently finely sampled observations. By the theory of quadratic variation, this same idea carries over to estimates of the integrated volatility over fixed horizons.

In the notation of ABDE, assume the logarithmic Nx1 vector price process, p_t, follows a multivariate continuous-time stochastic volatility diffusion,

$$dp_t = \mu_t dt + \Omega_t dW_t$$

Where W_t denotes a standard N-dimensional Brownian motion, and the process for the NxN positive definite diffusion matrix Ω_t is strictly stationary. Conditional on the sample path realization of μ_t and Ω_t the distribution of the continuously compounded h-period returns $r_{t+h,t} = p_{t+h} - p_t$ is then:

$$r_{t+h,t} \left| \sigma\left[\mu_{t+\tau}, \Omega_{t+\tau}\right]_{\tau=0}^{h} \sim N\left(\int_0^h \mu_{t+\tau} d\tau, \int_0^h \Omega_{t+\tau} d\tau\right)\right.$$

The integrated diffusion matrix is consequently a natural measure of the underlying h-period volatility. By the theory of quadratic variation, under weak regularity conditions,

$$\sum_j r_{t+j\Delta,\Delta} \bullet r'_{t+j\Delta,\Delta} - \int_0^h \Omega_{t+\tau} d\tau \rightarrow 0$$

almost surely for all t as the sampling frequency of returns increases ($\Delta \rightarrow 0$). Hence, ABDE show, by summing sufficiently finely-sampled high frequency returns, it is possible to construct ex-post realized volatility measures for the integrated latent volatilities that are asymptotically free of measurement error. This contrasts with the squared return over the forecast horizon which, while unbiased, is contaminated with measurement error which typically drowns out the predictable variation in the true latent volatility. Andersen and Bollerslev (1998) demonstrate the increased efficiencies afforded by higher frequency data in modeling two exchange rate volatility processes, for US$-DM and US$-JPY. The population coefficients of multiple determination, R^2, increase monotonically with sampling frequency from only 6.3% (DM) and 8.9% (JPY) for daily sampling, to in excess of 48% for both series with five minute sampling, while the measurement errors fall from 1.14 (DM) and 0.84 (JPY) for daily sampling to 0.004 or less for five minute sampling.

As volatility becomes effectively observable by these means, ABDE point out that standard statistical procedures can be used to analyze its distributional properties. This they proceed to do for the 30 component stocks within the Dow Jones Industrial Average (DJIA). In all cases they find that, while the daily returns display the characteristic negative skewness and high kurtosis, the standardized returns (using the realized volatility) are approximately normally distributed: the median kurtosis is reduced from 5.42 for raw returns to 3.13 for standardized returns. In parallel with this finding the authors that the log realized volatility is also approximately normal. This mirrors the findings of French, Schwert and Stambaugh (1987), who find that log monthly standard deviations constructed from the daily returns within the month is approximately Gaussian. Comparable findings are made by Andersen, Bollerslev, Diebold and Labys, 2000, (hereafter ABDL), with regard to the distribution of foreign exchange rate volatility.

These findings suggest that the unconditional distribution of daily returns should be well described by a continuous lognormal-normal mixture, as proposed by Clark (1973) in his Mixture-of-Distributions-Hypothesis.

Areal and Taylor, 2000, refine the realized volatility concept by introducing weighting schemes for computing realized volatility using intra-day volatility multipliers. Letting r_{ij}, $0 \le j \le n$ represent a set of n+1 intraday returns for day, intraday volatility is modeled in multiplicative form by:

$$\operatorname{var}(r_{ij} | \sigma_t) = \lambda_j \sigma_t^2 \qquad \text{with } \sum_{j=0}^{n} \lambda_j = 1$$

Where λ_j is the proportion of a trading day's total return variance, σ_t^2 that is attributed to period j. Here it is assumed that intraday returns are uncorrelated and that the multipliers are the same for all days t.
Under this scheme the realized variance for day t is estimated by weighting intraday squared returns as follows:

$$\sigma_t^2 = \sum_{j=0}^{n} \omega_j r_{tj}^2$$

To ensure conditionally unbiased estimates when intraday returns are uncorrelated, i.e $E\left[\hat{\sigma}_t^2 | \sigma_t^2\right] = \sigma_t^2$, it is necessary to apply the constraint $\sum_{j=0}^{n} \lambda_j \omega_j = 1$.

ABDE and related papers simply use $\omega_j = 1$, for all intraday periods j. Areal and Taylor show that in fact the variance of the estimator is minimized when

$$\omega_j = \frac{1}{(n+1)\lambda_j}$$

The Log Range as a Volatility Metric

The log range is defined as the difference between the highest and lowest log price between times iH and (i+1)H. The log volatility metric is then:

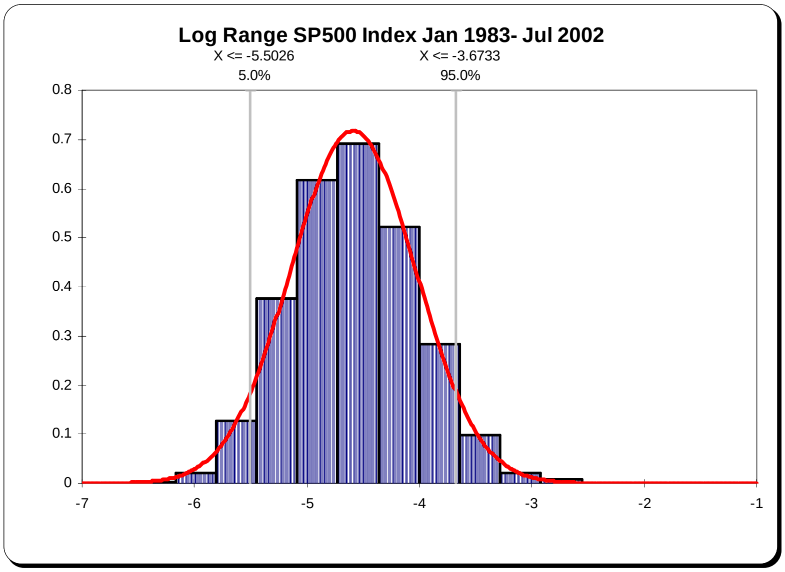
$$\ln\left| f\left(s_{iH,(i+1)H}\right)\right| = \ln\left[\sup_{iH<t<(1+1)H} s_t - \inf_{iH<t<(1+1)H} s_t\right]$$

$$= \ln\sigma_{iH} + \ln\left[\sup_{iH<t<(1+1)H} s^*_t - \inf_{iH<t<(1+1)H} s^*_t\right]$$

The log range is superior as a volatility metric on two counts. Firstly it is more efficient, in that the variance on the measurement errors is far smaller than for standard volatility metrics. Secondly, the distribution of the log range is very close to Normal, which makes it attractive for use in Gaussian QMLE models. To demonstrate the efficiency advantage of the log range, ABD apply a result due to Feller (1951) to derive its distribution, from which the superiority of the log range relative to the log absolute return is evident. Both metrics are linear functions of log volatility, but the standard deviation of the log range (0.29) is around one fourth that of the log absolute return (1.11).

The intuition behind the superior efficiency of the log range is simply that, on a volatile trading day when the market happens to close near the opening price, standard volatility measures will indicate low volatility, despite the large intraday fluctuations. The long range, on the other hand, will take account of the large intraday movement and correctly indicate a higher estimate of volatility. As ABD point out, the range has featured in a number of popular measures of market activity. Its superior performance characteristics have been discussed in academic literature since Parkinson (1980) and extensively researched by Garman and Klass (1983), Ball and Torous (1984), Rogers and Satchell (1991), and Yang and Zhang (2000).

The approximate normality of the log range is clearly illustrated in the figure below, which shows the distribution of the log range for the SP500 index. The higher skewness (0.24) and kurtosis (3.23) compare with the theoretical values (0.17 and 2.80 respectively), which are themselves close to those of a standard normal variate.



SP500 Index: Log Monthly Volatility

January 1983 – June 2002

In addition to the desirable properties of efficiency and normality, the authors demonstrate a third important property of the log range: robustness to certain types of market microstructure effects. To illustrate the robustness of the range to such effects ABD compare its properties to those of realized volatility, another highly efficient volatility proxy, in the presence of bid-ask bounce. In the presence of a bid-ask spread the observed price is a noisy version of the true price because it effectively equals the true price plus or minus half the spread. Because transactions tend to bounces between purchases and sales, the induced bid-ask bounce in observed prices increases the measured volatility of high-frequency returns. This effect is present in the case of high frequency returns and in the average size of squared high frequency returns. Because it entails summing the squared high frequency returns, each of which has an upward bias, the realized volatility contains a cumulated and therefore potentially large bias, which becomes more sever with sampling frequency. By contrast, the range is less likely to be seriously contaminated by bid-ask bounce. The observed daily maximum is likely to be at the ask and hence too high by half a spread, whereas the observed minimum is likely to be on ht e bid and hence too low by half a spread. On average, therefore, the range is inflated only by the average spread, which is small in most markets. The result is that, although in theory the log range is a less efficient estimator than realized volatility under ideal conditions, it may prove a more reliable metric under the usual market conditions.

The authors illustrate the principle as follows: suppose that the true log price s_t evolves as a random walk, $s_t = s_{t-1} + \mu_t$ with μ_t ~NID[0,σ^2_u]. Let the bid price be B_t = min[S_t – ticksize] and the ask price be A_t = min[S_t + ticksize], where S_t = exp(s_t) us the true price. We then take the observed price as $S(obs)_t = B_t q_t + A_t(1-q_t)$ where qt = Bernoullis[1/2]. The authors use as an example S0 = $25, ticksize = $1/16 and σ_u = 0.0011, which implies an annualized volatility of 30%, assuming 250 trading days a year. The population daily volatility is 1.87% and the realized volatility calculated using the true returns is close at 1.81%. By contrast the realized volatility based on the noisier observed returns is highly inflated at 6.70%. Market microstructure noise in the observed returns also affects the range-based volatility metric, but the effect is relatively minor. For true and observed prices the range based volatility estimates are 1.54% and 1.79% respectively.

The researchers perform a more systematic analysis using repeated samples with a variety of sampling frequencies (5-minute, 10-minute, 40-minute, 80 minutes, 3 hour, 6 hour and 20 hour), simulating one day of five-minute true and observed prices from a process with a volatility of true daily returns fixed at 1.87%. They then report the mean, standard deviations and RMSE of the realized and range-based volatility estimates. Considering first estimating volatility with true prices, the researchers find that realized volatility is unbiased regardless of the return interval, and its standard deviation decreases monotonically to zero as the return interval shrinks. In contrast, range-based volatility is downwards biased, regardless of the return interval, because the range on a discrete process must necessarily be less than for a continuous one. As the sampling interval gets shorter, this bias decreases monotonically, but the standard deviation increases monotonically. At five-minute intervals the efficiency (RMSE) of range-based volatility is between that of realized volatility computed using 3-hour and 6-hour returns, which accords with the results of Andersen and Bollerslev (1998). Overall, for a true price process, realized volatility dominates range-based volatility regardless of the sampling interval. The situation with an observed price process is, however, very different. The bid-ask bounce biases realized volatility upward, and the bias increases monotonically as the underlying return interval shrinks. To make matters worse the variability of realized volatility stays high as the return interval shrinks, because the benefits of using higher frequency data are outweighed by the negative effects of marker microstructure noise. As a results, the RMSE of realized volatility spikes sharply upward as the return interval shrinks. Bid-ask bounce affects range-based volatility differently. The discreteness of the process tends to bias the metric downwards but the bid-ask bounce tends to inflate it. The two biases trade off against once another, typically producing very good performance of range-based volatility in the presence of microstructure noise. Relative to realized volatility the range based estimator performs very well and its relative efficiency increases as the return interval narrows.

The researchers next proceed with a comparison of the performance characteristics of the log range vs. log absolute returns, using Monte-Carlo simulation. The results indicate that the log range is far more efficient as a volatility proxy. These advantages stem from two sources: the range is a much less noisy volatility measure and its Gaussian distribution property. To separate the two effects ABD compare the range-base quasi-maximum likelihood estimator to the exact maximum likelihood estimator for absolute returns. If the only benefit from using the range is its approximate normality the results for the quasi-maximum likelihood estimator should be similar to those for the exact maximum likelihood estimator for absolute returns. If, on the other hand, there is useful information about intraday volatility revealed by the log range estimator but not by absolute or squared returns, the sampling properties of the range-base quasi-maximum likelihood estimator should dominate the sampling properties of the exact maximum likelihood estimator for absolute returns.

The results of the research indicate that the majority, but not all, of the efficiency gain from using the log range is attributable to its distribution properties. In terms of bas the range-base quasi-maximum likelihood estimator and the exact maximum likelihood estimator for absolute returns perform equally well. However the RMSE's of the range-based parameter estimates are significantly smaller than for the corresponding exact maximum likelihood estimates. This demonstrates that the information about intraday volatility contained in the range plays and important role in the success of the range-based estimator.

Because the return-based estimators do not utilize intraday data, their sampling distributions are independent of the umber of trades per day N. This is not the case for the range-based estimator whose properties depend on the level of trading activity. In particular, where there are only a few trades per day, the observed range can be far from the true range of the underlying process and, as a result, can lead to a substantial mis-estimation of the true process volatility.

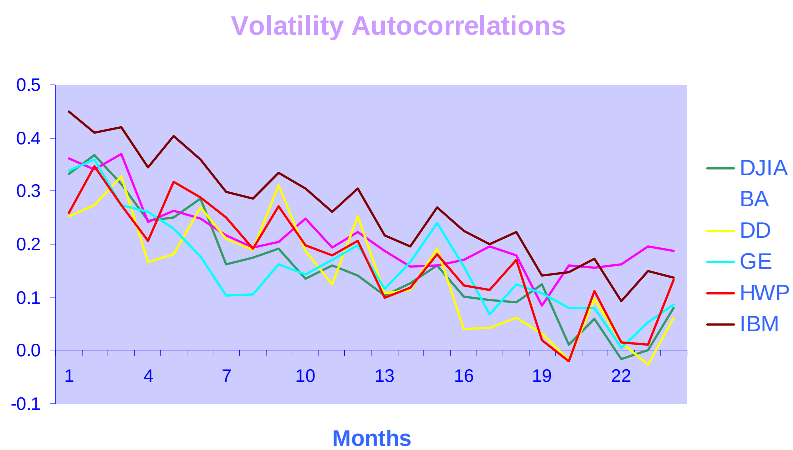
The authors examine the robustness of the range-based estimator to less frequent trading by comparing the results for N=100, 50 and 10. The general pattern is that as N decreases the range-based parameter estimates becoming increasingly biased. However, the performance of the range-based estimator clearly dominates the quasi-maximum likelihood estimator with log absolute returns, in terms of both bias and RMSE, for N=100 and N=50. Range based estimation is inferior to returns-only estimation only hen there are less than 10 trades a day.

Once the model has been estimated, the Kalman filter can be used to extract the latent stochastic volatility series. As the authors point out, the filter produces only linear projections, which coincide with conditional expectations only under the assumption of joint normality. Hence the extraction of latent volatilities is best unbiased when using a Gaussian volatility proxy, whereas the extraction is merely best linear unbiased for non-Gaussian proxies. This further confirms the superiority of the log range as a volatility estimator. Not only are the parameter estimates more accurate, but even for the same parameter values the efficiency of the log range as a volatility metric and the approximate normality of the log range yield more accurate volatility extractions.

Characteristics of Asset Volatility

Long Memory

The conditional distribution of asset volatility has been the subject of extensive empirical research in the last decade. The overwhelming preponderance of evidence points to the existence of pronounced long term dependence in volatility, characterized by slow decay rates in autocorrelations and significant correlations at long lags (e.g. Crato and de Lima, 1993, and Ding, Granger and Engle, 1993). ABDE find similar patterns for autocorrelations in the realized volatility processes for the Dow 30 stocks - autocorrelations remain systematically above the conventional Bartlett 95% confidence band as far out as 120 days. Comparable results are seen when autocorrelations are examined for daily log range volatility, as the figure below illustrates. Here we see significant autocorrelations in some stocks as far back as two years.



Volatility Autocorrelations for Selected DOW Stocks

Long Memory Detection and Estimation

Among the first to consider the possibility of persistent statistical dependence in financial time series was Mandelbrot (1971), who focused on asset returns. Subsequent empirical studies, for example by Greene and Fielitz (1977), Fama and French (1988), Porteba and Summers (1988) and Jegadeesh (1990), appeared to lend support for his findings of anomalous behavior in long-horizon stock returns. Tests for long range dependence were initially developed by Mandelbrot using a refined version of a test statistic, the Rescaled Range, initially developed by English hydrologist Harold Hurst (1951).

The classical rescaled range statistic is defined as

$$R/S(n) = \frac{1}{s_n}\left[\underset{1\le k\le n}{Max \sum_{j=1}^{k} (X_j - \overline{X}_n) - Min \sum_{j=1}^{k} (X_j - \overline{X}_n)} \right]$$

Where s_n the sample standard deviation:

$$s_n = \left[\frac{1}{n} \sum_j (X_j - \overline{X}_n)^2 \right]^{1/2}$$

The first term is the maximum of the partial sums of the first k deviations of X_j from the sample mean. Since he sum of all n deviations of the X_j's from their mean is zero, this term is always nonnegative. Correspondingly, the second term is always nonpositive and hence the difference between the two terms, known as the range for obvious reasons, is always nonnegative.

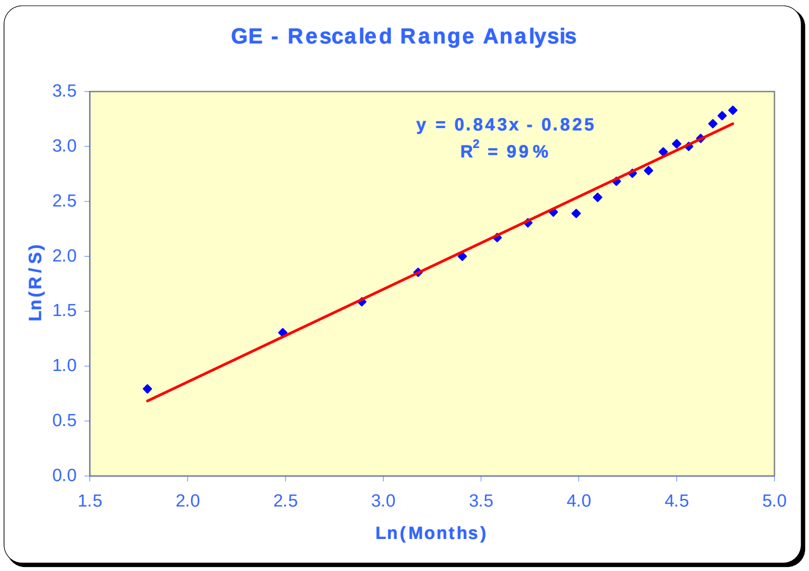
Madelbrot and Wallis (1969) use the R/S statistic to detect long range dependence in the following way. For a random process there is scaling relationship between the rescaled range and the number of observations n of the form:

$$R/S(n) \sim n^H$$

where H is known as the Hurst exponent. For a white noise process H = 0.5, whereas for a persistent, long memory process H > 0. The difference d = (H-0.5) represents the degree of fractional integration in the process.

Mandelbrot and Wallis suggest estimating the Hurst coefficient by plotting the logarithm of R/S(n) against log(n). For large n, the slope of such a plot should provide an estimate of H. The researchers demonstrate the robustness of the test by showing by Monte Carlo simulation that the R/S statistic can detect long-range dependence in highly non-Gaussian processes with large skewness and kurtosis. Madelbrot (1972) also argues that, unlike spectral analysis which detects periodic cycles, R/S analysis is capable of detecting nonperiodic cycles with periods equal to or greater than the sample period.

The technique is illustrated below for the volatility process of General Electric Corporation, a DOW Industrial Index component. The estimated Hurst exponent given by the slop of the regression, approximately 0.8, indicates the presence of a substantial degree of long-run persistence in the volatility process. Analysis of the volatility processes of other DOW components yield comparable Hurst exponent estimates in the region of 0.76 – 0.96.



Estimation of Hurst Exponent for GE Volatility Process

A major shortcoming of the rescaled range is its sensitivity to short-range dependence. Any departure from the predicted behavior of the R/S statistic under the null hypothesis need not be the result of long-range dependence, but may merely be a symptom of short-term memory. Lo (1991) show that this results from the limiting distribution of the rescaled range:

$$\frac{1}{\sqrt{n}} R/S(n) \Rightarrow V$$

Where V is the range of a Brownian bridge on the unit interval.

Suppose now that the underlying process $\{X_j\}$ is short range dependent, in the form of a stationary AR(1), i.e.

$$r_t = \rho r_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N(0,\sigma^2), \quad |\rho| \in (0,1)$$

The limiting distribution of R/S(n) / √n is V[(1+ρ)/(1-ρ)]$^{1/2}$. As Lo points out, for some common stocks the estimated autoregressive coefficient is as large at 0.5, implying that the mean of R/S(n) / √n may be biased upward by as much as 73%. In empirical tests, Davies and Harte (1987) show that even thought the Hurst coefficient of a stationary Gaussian AR(1) is precisely 0.5, the 5% Mandelbrot regression test rejects this null hypothesis 47% of the time for an autoregressive parameter of 0.3

To distinguish between long-range and short-term dependence, Lo proposes a modification of the R/S statistic to ensure that its statistical behavior is invariant over a general class of short memory processes, but deviates for long memory processes. His version of the R/S test statistic differs only in the denominator. Rather than using the sample standard deviation, Lo's formula applies the standard deviation of the partial sum, which includes not only the sums of squares of deviations for X_j, but also the weighted autocovariances (up to lag q):

$$\hat{\sigma}_n^2(q) = \frac{1}{n}\sum_{j=1}^{n}(X_j - \overline{X}_n)^2 + 2\sum_{j=1}^{q}\omega_j(q)\hat{\gamma}_j, \quad \omega_j(q) = 1 - \frac{j}{q+1}, \quad q < n$$

where the γ_j are the usual autocovariance estimators.

While in principle this adjustment to the R/S statistic ensures its robustness in the presence of short-term dependency, the problem remains of selecting an appropriate lag order q. Lo and MacKinlay (1989) have shown that when q becomes relatively large to the sample size n, the finite-sample distribution of the estimator can be radically different from its asymptotic limit. On the other hand, q cannot be taken too small as the omitted autocovariances beyond lag q may be substantial. Andrews (1991) provides some guidance on the choice of q, but since criteria are based on asymptotic behavior and little is known about the optimal choice of lag in finite samples.

Another method used to measure long-range dependence is the detrended fluctuation analysis (DFA) approach of Peng et al(1994) and further developed by Viswanathan et al (1997). Its advantage over the rescaled range methodology is that it avoids the spurious detection of apparent long-run correlation due o non-stationarities. In the DFA approach the integrate time series y(t') is obtained:

$$y(t') = \sum_{T=1}^{t'} x(t).$$

The series y(t') is divided into non-overlapping intervals each containing m data points and a least squares line is fitted to the data. Next, the root mean square fluctuation of the detrended time series is calculated for all intervals:

$$F(m) = \sqrt{\frac{1}{T}\sum_{t'=1}^{T}\left[y(t') - y_m(t')\right]^2}$$

A log-log plot of F(m) vs the interval size m indicates the existence of a power-scaling law. If there is no correlation, or only short term correlation, then F(m) $\propto$ $m^{1/2}$, but if there is long-term correlation then F(m) will scale at rates greater than ½.

A third approach is a semi-parametric procedure to obtain an estimate of the fractional differenceing parameter d. This techniques, due to Geweke and Porter-Hudak (GPH), is based on the slope of the spectral density around the angular frequency w = 0. The spectral regression is defined by:

$$\ln\{I(\omega_\lambda\} = a + b\ln\left\{4\sin^2\frac{\omega_\lambda}{2}\right\} + n_\lambda, \quad \lambda = 1,...,\upsilon$$

Where $I(w_\lambda)$ is the periodogram of the time series at frequencies $w_\lambda = 2\pi\lambda/T$ with l = 1, . . ., (T-1)/2. T is the number of observations and υ is the number of Fourier frequencies included in the spectral regression. The least squares estimate of the slope of the regression line provides an estimate of d. The error variance is $\pi^2/6$ and allows for the consecution of the t-statistics for the fractional differencing parameter d. An issue with this procedure is the choice of υ, which is typically set to T1/2, with Sowell (1992) arguing that u should be based on the shortest cycle associated with long-run correlation.

The final method we consider is due to Sowell (1992) and is a procedure for estimating stationary ARFIMA models of the form:

$$\Phi(L)(1-L)^d(y_t-\mu)=\Theta(L)\varepsilon_t$$

Where Φ and Θ are lag polynomials, d is the fractional differencing parameter, μ is the mean of the process $y_t \sim N(\mu, \Sigma)$ and ε_t is an error process with zero mean and constant variance σ_e^2. We can use any set of exogenous regressors to explain the mean: $z = \mathbf{y} - \mu$, $\mu = f(\mathbf{X},\beta)$.

The spectral density function is written in terms of the model parameter d, from which Sowell derives the autocovariance function at lag k in the form:

$$\gamma(k)=\frac{1}{2\pi}\int_0^{2\pi} f(W)e^{iwk}dw$$

The parameters of the model are then estimated by exact maximum likelihood, with log likelihood:

$$\log L(d,\phi,\theta,\beta,\sigma_e^2)=-\frac{T}{2}\log(2\pi)-\frac{1}{2}\log|\mathbf{\Sigma}|-\frac{1}{2}\mathbf{z}'\mathbf{\Sigma}^{-1}\mathbf{z}$$

Structural Breaks

Granger and Hyung, 1999, take a different approach to the analysis of long term serial autocorrelation effects. Their starting point is the standard I(d) representation of an fractionally integrated process y_t of the form:

$$(1-L)^d y_t = \varepsilon_t$$

where d is the fractional integration parameter and, from its Maclaurin expension

$$(1-L)^d = \sum_{j=0}^{\infty}\pi_j L^j, \quad \pi_j = \frac{j-1-d}{j}\pi_{j-1}, \quad \pi_0 = 1$$

The researchers examine the evidence for structural change in the series of absolute returns for the SP500 Index by applying the sequential break point estimation methodology of Bai (1997) and Bai and Perron (1998) and Iterative Cumulative Sums of Squares (ICSS) technique of Aggarwal, Inclan and Leal, 1999. Bai's procedure works as follows. When the break point is found at period k, the whole sample is divided into two subsamples with the first subsample consisting of k observations and the second containing the remaining (T-k) observations. A break point is then estimated for the subsample where a hypothesis test of parameter consistency is rejected. The corresponding subsample is then divided into further subsamples at the estimated break point and a parameter constancy test performed for the hierarchical subsamples. The 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. Bai shows how the sequential procedure coupled with hypothesis testing can yield a consistent estimate for the true number of breaks.

Aggarwal, Inclan and Leal's (1999) approach uses the Iterative Cumulative Sums of Squares (ICSS) as follows. We let $\{e_t\}$ denote a series of independent observations from a normal distribution with zero mean and unconditional variance σ_t^2. The variance within each interval is denoted by τ^2_j, j = 0,1, . . . ,N_t, where N_t is the total number of variance changes in T observations and $1 < k_1 < k_2 < . . . < k_{NT} < T$ are the set of change points.

So

$$\sigma_t = \tau_j \quad k_j < t < k_{j+1}$$

To estimate the number of changes in variance and the point in time of the shift a cumulative sum of squares is used.

Let

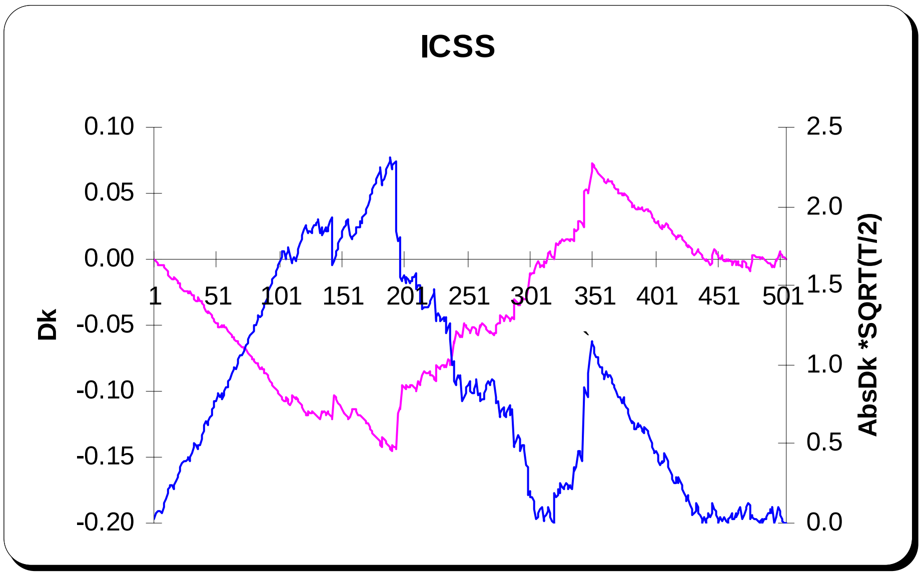
$$C_k = \sum_{t=1}^{k} \varepsilon_t^2, \quad k = 1,...,T$$

be the cumulative sum of the squared observations from the start of the series until the kth point in time. Then define Dk = $(C_k / C_T) - k/T$.

If there are no changes in variance over the sample period, the D_k oscillate around zero. Critical values based on the distribution of Dk under the null hypothesis of no change in variance provide upper and lower bounds to detect a significant change in variance with a known level of probability. Specifically, if maxk$\sqrt{(T/2)}$| D_k| exceeds

1.36, the 95th percentile of the asymptotic distribution, then we take k*, the value of k at which the maximum value is attained as an estimate of the change point.

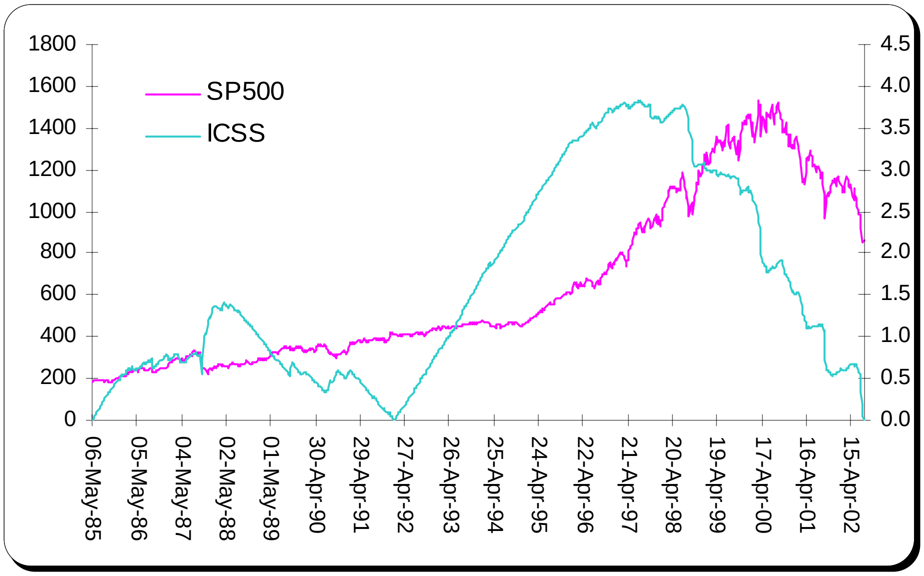
The figure below illustrates the procedure for a simulated GBM process with initial volatility of 20%, which changes to 30% after 190 periods, and then reverts to 20% once again in period 350. The test statistic $\sqrt{(T/2)}|Dk|$ reaches local maxima at t = 189 (2.313) and t = 349 (1.155), clearly and accurately identifying the two break points in the series.



Testing for Structural Breaks in Simulated GMB Process Using Iterative Cumulative Sums of Squares

A similar analysis is carried out for the series of weekly returns in the SP500 index from April 1985 to April 2002. Several structural shifts in the volatility process are apparent, including the week of 19 Oct 1987, 20 July 1990 (Gulf War), the market tops around Aug 1997, Aug 1998 and Oct 2000.

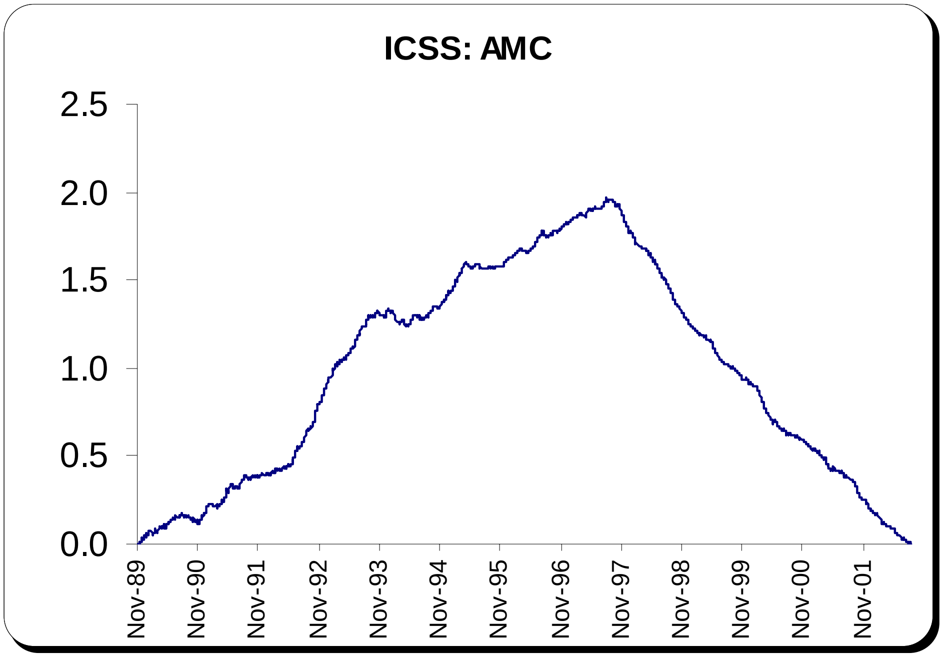
In their comprehensive analysis of several emerging and developed markets, Aggarwal et al identify numerous structural shifts relating to market crashes, currency crises, hyperinflation and government intervention, including, to take one example, as many as seven significant volatility shifts in Argentina over the period from 1985 – 1995.



Testing for Structural Breaks in SP500 Index Returns
Using Iterative Cumulative Sums of Squares

It is common for structural breaks to result in ill-conditioning in the volatility processes distribution, often in the form of excess kurtosis. This kind of problem can sometimes be resolved by modeling the different regime segments individually. Less commonly, regime shifts can produce spurious long memory effects. For example, Granger and Hyung estimate the degree of fractional integration d in daily SP500 returns for 10 subperiods from 1928 – 1991 using the standard Geweke and Porter-Hudak approach. All of the subperiods have strong evidence of long memory in the absolute stock return. They find clear evidence of a positive relationship between the time-varying property of d and the number of breaks, and conclude that the SP500 Index absolute returns series is more likely to show the “long memory” property because of the presence of a number of structural breaks in the series rather than being an I(d) process.
Stocks in Asian-Pacific markets typically exhibit volatility regime shifts at around the time of the regional financial crisis in the latter half of 1997. The case of the ASX200 Index component stock AMC is typical (see figure below). Rescaled range analysis of the entire volatility process history leads to estimates of fractional integration of the order of 0.2. But there is no evidence of volatility persistence is the series post-

1997. The conclusion is that, in this case, apparent long memory effects are probably the result of a fundamental shift the volatility process.



Structural Breaks in the Asian Crisis Period for ASX Component Stock AMC.

Volatility Models

Continuous Time Stochastic Volatility Models

We follow the terminology of Alizahed, Brandt and Dielbold (2001) (hereafter ABD) in describing the evolution of an asset price S as a diffusion process with instantaneous drift μ and volatility σ. Both the drift and volatility depend on a latent state variable v, which itself evolves as a diffusion:

$$
\begin{aligned}
dS_t &= \mu(S_t, v_t) + \sigma(S_t, v_t) dW_{St} \\
dv_t &= \alpha(S_t, v_t) + \beta(S_t, v_t) dW_{vt}
\end{aligned}
$$

where W_{St} and W_{vt} are two Weiner processes with correlation $dW_{St}dW_{vt} = \theta(S_t, v_t)dt$.

ABD focus on the following streamlined version:

$$
\begin{aligned}
\frac{dS_t}{S_t} &= \mu dt + \sigma_t dW_{St} \\
d \ln \sigma_t &= \alpha(\ln \overline{\sigma} - \ln \sigma_t) dt + \beta dW_{vt}
\end{aligned}
$$

Here the log volatility $\ln\sigma$ of returns dS/S is the latent state variable, which evolves as mean-reverting Ornstein-Uhlenbeck process with man reversion parameter α.

In the decretized version of the above we rely on N discrete time price realizations, drawn from a sample period T with intervals T/N, to draw inferences about the continuous-time model. Within each interval I, i.e. between times iH and (i+1)H, volatility is assumed to be constant at $\sigma_t = \sigma_{ih}$, but from one interval to the next, volatility is stochastic. The security price evolves as geometric Brownian motion:

$$
\frac{dS_t}{S_t} = \mu dt + \sigma_{iH} dW_{St}
$$

while the discretized log volatility has a conditional Gaussian distribution of approximately:

$$
N\left[\ln \overline{\sigma} + (1 - \alpha H)(\ln \sigma_{iH} - \ln \overline{\sigma}), \beta^2 H\right]
$$

This formulation can readily be extended to two-factors in the style of Engle and Lee (1999), with transition equations:

$$\ln \sigma_{(i+1)H} = \ln \overline{\sigma} + \ln \sigma_{1,(i+1)H} + \ln \sigma_{2,(i+1)H}$$

with

$$\ln \sigma_{1,(i+1)H} = \rho_{1,H} \ln \sigma_{1,iH} + \beta_1 \sqrt{H} \nu_{1,(i+1)H}$$

$$\ln \sigma_{2,(i+1)H} = \rho_{2,H} \ln \sigma_{2,iH} + \beta_2 \sqrt{H} \nu_{2,(i+1)H}$$

and where the volatility innovations v_1 and v_2 are contemporaneously and serially independent N[0.1] variates.

ABD apply two-factor models in this form to five US Dollar exchange rate series. They find that one factor has highly persistent dynamics while the other has transient dynamics, with each accounting for approximately half of the long-run (unconditional) variance of log volatility. As the authors report, the result is intuitively appealing and in line with properties of volatilities estimated using very different procedures, such as the realized volatilities of Andersen, Bollerslev, Diebold and Labys (2001), which appear to show persistent movement, with high frequency noise superimposed.

Range-Based Volatility Models

Alizadeh, Brandt, and Diebold (2001) follow a related, but slightly different track. They assume that the log stock price s follows a drift-less Brownian motion dσ = σdW. The volatility of daily log returns, denoted h= σ/√252, is assumed constant within each day, at ht from the beginning to the end of day t, but is allowed to change from one day to the next, from ht at the end of day t to ht+1 at the beginning of day t+1.6 Under these assumptions, the researchers show that the log range, defined as:

$$D_t = \ln\left(\max_{\tau \in [t,t+1]} s_\tau - \min_{\tau \in [t,t+1]} s_\tau \right)$$

is to a very good approximation distributed as

$$D_t \sim N\left[0.43 + \ln h_t, 0.29^2\right]$$

where N[*m; v*] denotes a Gaussian distribution with mean *m* and variance *v*. The above equation demonstrates that the log range is a *noisy* linear proxy of *log* volatility ln *ht*. By contrast, the log absolute return has a mean of 0.64 + ln *ht* and a variance of 1*:*11. However, the distribution of the log absolute return is far from Gaussian. The fact that both the log range and the log absolute return are linear log volatility proxies (with the same loading of one), but that the standard deviation of the log range is about one-quarter of the standard deviation of the log absolute return, makes clear that the range is a much more informative volatility proxy. It also makes sense of the finding of Andersen and Bollerslev (1998) that the daily range has approximately the same informational content has sampling intra-daily returns every four hours.

It is well known that the range suffers from a discretization bias because the highest (lowest) stock price observed at discrete points in time is likely to be lower (higher) than the true maximum (minimum) of the underlying diffusion process. It follows that the observed range is a downward-biased estimate of the true range (which in turn is a noisy proxy of volatility). Rogers and Satchell (1991) devise a correction of the observed range that virtually eliminates this bias. However in many cases the discretization bias is not likely to be a problem where the underlying asset is very liquid and the time that elapses between trades (recorded prices) is negligible.

Except for the model of Chou (2001), GARCH-type volatility models rely on squared or absolute returns (which have the same information content) to capture variation in the conditional volatility *ht*. Since the range is a more informative volatility proxy, it makes sense to consider range-based GARCH models, in which the range is used in place of squared or absolute returns to capture variation in the conditional volatility. This is particularly true for the EGARCH framework of Nelson (1990), which describes the dynamics of log volatility (of which the log range is a linear proxy).

The researchers consider variants of the EGARCH framework introduced by Nelson (1990). In general, an EGARCH(1,1) model performs comparably to the GARCH(1,1) model of Bollerslev (1987). However, for stock indices the in-sample evidence reported by Hentschel (1995) and the forecasting performance presented by Pagan and Schwert (1990) show a slight superiority of the EGARCH specification. One reason for this superiority is that EGARCH models can accommodate asymmetric volatility (often called the "leverage effect," which refers to one of the explanations of asymmetric volatility), where increases in volatility are associated more often with large negative returns than with equally large positive returns.

The one-factor rang-based model (REGARCH 1) takes the form:

$$D_t \sim N\left[0.43 + \ln h_t, 0.29^2\right]$$

$$\ln h_t - \ln h_{t-1} = k_h(\vartheta - \ln h_{t-1}) + \phi_h X_{t-1}^D + \delta_h R_{t-1} / h_{t-1}$$

where the returns process R_t is conditionally Gaussian: $R_t \sim N[0, h_t^2]$
and the process innovation is defined as the standardized deviation of the log range from its expected value:

$$X_{t-1}^D = (D_{t-1} - 0.43 - \ln h_{t-1}) / 0.29$$

Following Engle and Lee (1999), the researchers also consider multi-factor volatility models. In particular, for a two-factor range-based EGARCH model (REGARCH2), the conditional volatility dynamics) are as follows:

$$\ln h_t - \ln h_{t-1} = k_h(\ln q_{t-1} - \ln h_{t-1}) + \phi_h X_{t-1}^D + \delta_h R_{t-1} / h_{t-1}$$
$$\ln q_t - \ln q_{t-1} = k_q(\vartheta - \ln q_{t-1}) + \phi_q X_{t-1}^D + \delta_q R_{t-1} / h_{t-1}$$

where ln *qt* can be interpreted as a slowly-moving stochastic mean around which log volatility ln *ht* makes large but transient deviations (with a process determined by the parameters k_h, ϕ_h and δ_h).
The parameters θ, k_q, ϕ_q and δ_q determine the long-run mean, sensitivity of the long run mean to lagged absolute returns, and the asymmetry of absolute return sensitivity respectively.

The intuition is that when the lagged absolute return is large (small) relative to the lagged level of volatility, volatility is likely to have experienced a positive (negative) innovation. Unfortunately, the absolute return is a rather noisy proxy of volatility, suggesting that a substantial part of the volatility variation in GARCH-type models is driven by *proxy noise* as opposed to true information about volatility. In other words, the noise in the volatility proxy introduces noise in the implied volatility process. In a volatility forecasting context, this noise in the implied volatility process deteriorates the quality of the forecasts through less precise parameter estimates and, more importantly, through less precise estimates of the current level of volatility to which the forecasts are anchored.

ABD test a selection of single and multi-factor returns-based and range-based models using daily data on the SP500 index over the period from January 1983 to September 2001. Perhaps unsurprisingly, they find that range-based models explain ranges better, while returns-based models are better at explaining squared returns. A more interesting finding is that two-factor models with some asymmetry are generally favored over their one-factor or symmetric equivalents. The model selection criteria exhibit greater variation across different specifications when they are computed with range data rather than returns data. This means that the information contained in the range allows us to distinguish between competing models that are virtually indistinguishable based n returns. The researchers also test the models for evidence of lack of fit using Engle's ARCH-LM test and find that here too the range-based tests are substantially more powerful than the return-based test. The tests reinforce the importance of both multiple factors and at least partial asymmetries. The researchers also examine the models for sign bias and find that only the range-based two-factor models with asymmetries pass all of the specification tests. The evidence for the superiority of two factor models in the entire-sample fit carries over into out-of-sample tests also. In almost every case the R-squared of the range-based forecasts is higher than that pf the corresponding return-based forecast. There is also consistent evidence in the results of the, admittedly modest, superiority of the two-factor models. The performance of the two-factor models is quite impressive even at longer time horizons, with forecast R-squares in the region of 0.2. This contradicts the usual perception and related empirical evidence of West and Cho (1995) and Christoffersen and Diebold (2000) that volatility predictability is a short-horizon phenomenon.

Volatility Feedback Models

It seems plausible that the considerable fluctuations which characterize volatility process in general may have some significant impact on asset returns and process. One interesting approach to modeling the interaction between asset volatility and asset returns processes is the volatility feedback models developed by Campbell and Hentschel, 1992, amongst others. Volatility feedback is an appealing idea because it provides credible explanations of some of the stylized facts about asset returns. For example, large negative returns arise more frequently than large positive returns, so returns distributions are negatively skewed. In similar vein, Bollerslev, 1987, has shown that excess kurtosis in returns process cannot simply be the result of changing volatility, because it is still present after the returns processes have been normalized by their estimated conditional volatility. Finally, stock returns tend to be negatively correlated with future volatility. This asymmetry effect was identified by Black, 1976, who argued that it could be due to the increase in leverage that occurs when stock process decline. However, studies by Christie, 1982, and Schwert, 1989, indicate that asymmetry effects are to large to be fully accounted for by leverage alone.

As Campell and Hentschel argue, volatility feedback provides a means of explaining these characteristics of returns, even if the underlying shocks to the processes are Gaussian. Suppose there is a large piece of good news about future dividends. Because volatility is persistent, large pieces of good news tend to be followed by other large pieces of news, so the initial news about future dividends will increase future expected volatility. This in turn will increase the required return on stock and lower the stock price, dampening the positive impact of the dividend news. Next, consider the impact of a large piece of negative news about future dividends. Here too the stock price falls because the higher volatility increases the required return on the stock and lowers the stock price. But now the volatility effect amplifies the negative impact of the dividend news. Large negative stock returns are therefore more common than large positive returns, and the amplification of negative returns can produce excess kurtosis. By contrasts, small pieces of good news, or even no news at all, permit stock prices to rise because “no news is good news” about future volatility. Volatility feedback therefore implies that stock price movements are correlated with future volatility.

The authors review previous research in the area. Brown, Harlow and Tinic, 1988, show that stock price reactions to unfavorable news tend to be larger than reactions to favorable events, a finding that they attribute to volatility feedback. Porteba and Summers, 1986, argue on the other hand hat volatility feedback cannot provide an adequate explanation because changes in volatility are too short-lived. French, Schwert and Stambaugh, 1987, regress stock returns on innovations in volatility and find a negative coefficient, which they attribute to a volatility feedback effect. Haugen Talmor and Torous, 1991, produce similar findings. Early research tended to discuss volatility feedback effects informally. Campbell and Hentschel provide a formal model for the process. To do so, they choose a form of GARCH model which enables returns to be correlated with future volatility, the QGARCH (quadratic GARCH) model of Engle, 1990, and Sentana, 1991. Their preference for the QGARCH model is explained by the fact that it is analytically tractable and captures the phenomenon of predictive asymmetry without requiring volatility feedback per se. As the authors explain, this is desirable since it appears that leverage effects account for at least some of the predictive asymmetry. Since the basic QGARCH model does not fit the negative skewness and excess kurtosis of returns, the researchers extend the model in a way that amplifies the volatility feedback effect and so permits these characteristics to arise. They begin by defining the one-period natural log real holding return on a stock as:

$$h_{t+1} \equiv \log(P_{t+1} + D_{t+1}) - \log(P_t)$$

where P_t is the real stock price at the end of period t (ex dividend) and D_t is the real dividend paid during period t. A Taylor series expansion gives:

$$h_{t+1} \approx k + \rho p_{t+1} + (1-\rho)d_{t+1} - p_t$$

where lower case letters are used for logs. The parameter ρ is the average ratio of the stock price to the sum of the stock price and the dividend, a number slightly smaller than 1, and the constant k is a nonlinear function of ρ. The equation holds ex-post, but it also holds ex-ante as an expected difference equation. Further manipulation leads to the following expression for the unexpected stock return:

$$h_{t+1} - E_t h_{t+1} = (E_{t+1} - E_t)\sum_{j=0}^{\infty} \rho^j \Delta d_{t+1+j} - (E_{t+1} - E_t)\sum_{j=0}^{\infty} \rho^j \Delta h_{t+1+j}$$

or, more succinctly, $\upsilon_{h,t+1} = \eta_{d,t+1} - \eta_{h,t+1}$

where $\upsilon_{h,t+1}$ denotes the unexpected stock return at time t+1, and the $\eta_{d,t+1}$ and $\eta_{h,t+1}$ denote news about dividends and future returns respectively.

Hence, as this relationship makes clear, if the unexpected return is negative, then either expected future dividend growth must be lower, or expected future stock returns must be higher, or both. The researchers treat the news about dividends component $\eta_{d,t+1}$ as an exogenous shock which follows a conditionally normal quadratic GARCH (QGARCH) process. The QGARCH(1,1) model is:

$$\eta_{d,t+1} \sim N(0, \sigma_t^2)$$

with $$\sigma_t^2 = \omega + \alpha(\eta_{d,t} - b)^2 + \beta\sigma_{t-1}^2$$

The parameters ω, α and β must al be positive to ensure that the conditional variance is non-negative. The parameter α measures the proportions of today's squared returns that contributes to future volatility, while the quantity $(\alpha + \beta)$ measures the persistence of volatility. The unconditional variance of the process is

$$(\omega + \alpha\beta^2)/(1 - (\alpha + \beta))$$

When the parameter b =0, the QGARCH model reduces to a standard GARCH model. A positive value of b introduces a negative correlation between the dividend news $\eta_{d,t}$ and the conditional volatility of the dividend news in the next period, s_t^2, because a negative return will increase volatility more than a positive return of the same size. Hence the QGARCH model allows us to model volatility asymmetry effects.

Turning to the other determinant of the stock return, news about future expected returns, the authors assume that the conditional expected return $E_t h_{t+1}$ is determined by the volatility of the news variable $\eta_{d,t+1}$:

$$E_t h_{t+1} = \mu + \gamma E_t \eta_{d,t+1}^2 = \mu + \gamma\sigma_t^2$$

Here γ is usually interpreted as the coefficient of relative risk aversion.

Using the above equations the authors show that the expected return at any future date can be written as:

$$E_t h_{t+1+j} = \mu + \gamma \frac{\omega + \alpha b^2}{1 - (\alpha + \beta)} + \gamma(\alpha + \beta)^j \left(\sigma_t^2 - \frac{\omega + \alpha b^2}{1 - (\alpha + \beta)} \right)$$

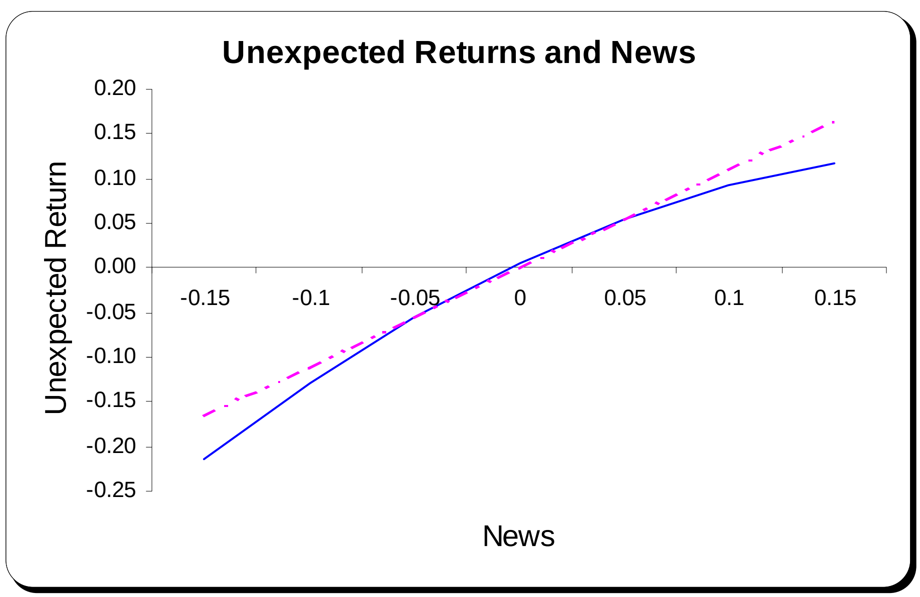
The second term on the RHS of this equation is gamma times the unconditional variance of the news process. The third term is gamma time the deviation of today's conditional variance from the unconditional variance, discounted using the persistence of volatility $\alpha + \beta$. Using this equation, the researchers are able to show that the stock return can be expressed as follows:

$$h_{t+1} = \mu + \gamma\sigma_t^2 + \kappa\eta_{d,t+1} - \lambda\left(\eta_{d,t+1}^2 - \sigma_t^2\right)$$

where $\kappa = 1+2\lambda b$ and $\lambda = \dfrac{\gamma\rho\alpha}{1-\rho(\alpha+\beta)}$

The first three terms of this equation describe the standard GARCH-M model. The final term says that an unusually large realization of dividend news will increase volatility, lower the stock price and cause a negative unexpected stock return. On the other hand a positive piece of dividend news will tend to reduce volatility and increases the stock return. The strength of the volatility feedback is measured by the parameter λ.

The researchers apply this model to explore the characteristics of returns processes. The chart below, which is similar to one produced by the authors, shows the relationships between dividend news $\eta_{d,t+1}$ and the unexpected stock return $\upsilon_{h,t+1} = \kappa\eta_{d,t+1} - \lambda\left(\eta_{d,t+1}^2 - \sigma_t^2\right)$ fusing parameter values λ = 2.398 and k = 1.012, estimated from postwar daily US data. The unexpected stock return (the solid line) lies above the 45 degree line where the absolute value of the news is less than its conditional standard deviation. This is the "no news is good news effect". If there is no dividend news at all, the stock market rises because the absence of dividend news implies that volatility and required returns will tend to be lower in future. Conversely the unexpected return lies below the 45-degree line at the left and right of the figure, where dividend news is large in absolute terms. Large declines in stock prices are amplified, while large increases are dampened. In fact the maximum return is achieved when $\eta_{d,t+1} = \kappa / 2\lambda$. Any larger piece of good dividend news actually lowers the stock return because the indirect volatility effect outweighs the direct dividend effect. The degree of curvature in the relation between news and returns depends on the level of σ_t^2. If σ_t^2 is small, then the quadratic term has little weight relative to the linear term in the relationship. On the other hand, if it is large, the quadratic term becomes much more important. The authors that that in fact the distribution of unexpected returns, standardized by σ_t, is a mixture of a normal distributions and a demeaned, negative Chi-Squared distributions with one degree of freedom. The normal distribution has weight $\kappa / (\kappa + \lambda\sigma_t)$ while the negative Ch-Squared distribution has weight $\lambda\sigma_t$ / $(\kappa + \lambda\sigma_t)$. In times of low volatility, returns are very close to normal, but in times of high volatility returns take on some of the characteristics of the negative Chi-Squared distribution. This shifting of distribution of returns is the result of a very important characteristic of GARCH and QGARCH models: the volatility of variance increases very rapidly with the level of variance. The researchers show that the conditional variance σ_{t+1}^2 is proportional to σ_t^4. This results in the volatility clustering phenomenon, where long periods of relative calm are interrupted by episodes of high and rapidly changing volatility. It also means that volatility feedback distorts the distribution of returns away from the normal more strongly when volatility is high than when it is low. As volatility increases, the variance of news about dividends increases with σ_t^2, but the variance of news about variance increases with σ_t^4. As the researchers point out, the key point is not that the variance of GARCH variance increases with its level – a common property of processes constrained to be non-negative – but that the variance of GARCH variance increases more than proportionally with its level.



Unexpected Returns and News

The researchers explore some of the other properties of the returns process in this model. Deriving an expression for skewness they show that the returns process is negatively skewed and the skewness increases with the conditional variance.

Further, the conditional distribution of stock returns has fat tails, even though the GARCH process for dividend news is conditionally normal. The conditional excess kurtosis increases with $\lambda \sigma_t^2$ to a limit of 12, the excess kurtosis of a Chi-Squared distribution with 1 degree of freedom. The model also generates a form of predictive asymmetry. There are two sources of correlation between return sand future volatility. The first is the QGARCH effect: as σ_t^2 increases this correlation approaches zero, but when σ_t^2 declines, it approaches +/- 1. The second source of correlation between returns and future volatility is the volatility feedback effect: this approaches –1 as volatility increases but approaches zero as volatility declines. The reason for this behavior is that news about variance becomes more important relative to the underlying news as the level of variance increases.

The researchers apply their model to monthly and daily data on excess stock returns in US stocks over the period from 1926-88. They find hat volatility persistence is very close to one, they are able to strongly reject the hypothesis that it equals one. The parameter estimates imply that volatility shocks have a half-life of between 12 and 18 months in monthly data and around six months in the daily data. Likelihood ratio tests strongly reject the GARCH model in favor of the QGRACH model, which appears able to capture most of the predictive asymmetry in the data. The λ coefficient is highly significant, emphasizing the statistical importance of volatility feedback effects. The authors explore the economic importance of the effect by examining the volatility discount, the log difference between the actual stock price and the price that would prevail in the absence of uncertainty about future dividends. Their results imply a volatility discount that is normally around 10%, rising above 12% in the 1930's and almost 14% in October 1987. The impact of volatility feedback on the variance of returns is found to be small, but it has a more important effect on the skewness and excess kurtosis of returns. Here it is found that feedback effects explain somewhat less than half the skewness and excess kurtosis. The implication appears to be that changing volatility generally has little effect on the level of stock prices. But during periods of high volatility, the feedback effect can become dramatically more important.

References

Abadir, K.M., Taylor, A.M.R., 1999. On the definitions of co- integration. Journal of Time Series Analysis 20, 129–137.

Aggarwal, R., Inclan, C., Leal, R., 1999, Volatility in Emerging Stock Markets, Journal of Financial and Quantitative Analysis, Forthcoming.

Andersen, T.G., and Bollerslev, T. (1997), “Heterogeneous Information Arrivals and Return Volatility Dynamics: Uncovering the Long-Run in High Frequency Returns,” *Journal of Financ*e, 52, 975-1005.

Andersen, T.G., and Bollerslev, T. (1998), “Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts,” *International Economic Revie*w, 39, 885-905.

Andersen, T.G., Bollerslev, T., Diebold, F.X., and Labys, P. (1999), “The Distribution of Exchange Rate Volatility,” Wharton Financial Institutions Center Working Paper 99-08 and NBER Working Paper 6961.

Andersen, T.G., and Sorensen, B.E. (1997), “GMM and QML Asymptotic Standard Deviations in Stochastic Volatility Models,” *Journal of Econometric*s, 76, 397-403.

Andrews, D, 1991, “Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation”, Economterica, 59, 817-858.

Alizadeh, S., M.W. Brandt, and F.X. Diebold, 2001, Range-based estimation of stochastic volatility models, *Journal of Finance*, forthcoming.

Bai, J., 1997, Estimating Multiple Breaks One at a Time, Econometric Theory, 13, 315-352.

Bai, J., Perron, P., 1998, Estimating and Testing Linear Models with Multiple Structural Changes, Econometrica, 66, 47-78.

Baillie, R.T., 1996. Long memory processes and fractional integration in econometrics. Journal of Econometrics 73, 5–59.

Baillie, R.T., Bollerslev, T., 1994a. Cointegration, fractional cointegration and exchange rate dynamics, Journal of Finance 49, 737–745.

Baillie, R.T., Bollerslev, T., 1994b. The long memory of forward premium. Journal of International Money and Finance 13, 565–571.

Baillie, T.R., Bollerslev, T., Mikkelsen, H.O., 1996. Fractionally integrated generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 74, 3–30.

Ball, C.A., and Torous, W. (1984), “The Maximum Likelihood Estimation of Security Price Volatility: Theory, Evidence, and Application to Option Pricing,” *Journal of Business*, 57, 97-112.
Beckers, S. (1983), “Variance of Security Price Returns Based on High, Low, and Closing Prices,” *Journal of Business*, 56, 97-112.
Black, F., and Scholes, M. (1973), “The Pricing of Options and Corporate Liabilities,” *Journal of Political Economy*, 81, 637-654.
Bollerslev, T., Chou, R.Y., and Kroner, K.F. (1992), “ARCH Modeling in Finance: A Selective Review of the Theory and Empirical Evidence,” *Journal of Econometrics*, 52, 5-59.
Bollerslev, T., Mikkelsen, H.O., 1996. Modelling and pricing long memory in stock market volatility, Journal of Economics 73, 151–184.
Brandt, M. and Jones, C., 2002, Volatility Forecasting with Range-Based EGARCH Models, forthcoming.
Breidt, J.F., Crato, N., de Lima, P., 1998. The detection and estimation of long memory in stochastic volatility. Journal of Economics 83, 325–348.
Campbell, J. and Hentschel, L., 1992, No News is Good News: An Asymmetric Model of Changing Volatility in Stock Returns, NBER working paper 3742.
Chou, R.Y., 2001, Forecasting financial volatilities with extreme values: The conditional autoregressive range (CARR) model, Working Paper, Adademia Sinica.
Christie, A., 1982, The Stochastic Behavior of Common Stock Variances: Value, Leverage and Interest Rate Effects, Journal of Financial Economics 10, 407-432
Danielsson, J. (1994), “Stochastic Volatility in Asset Prices: Estimation with Simulated Maximum Likelihood,” *Journal of Econometrics* 64, 375-400.
Davies, R and Harte, D, 1987, “Tests for Hurst Effect”, Biometrika, 74, 95-101.
Ding, Z., Granger, C.W.J., 1996. Modelling volatility persistence of speculative returns: a new approach. Journal of Econometrics 73, 185–215.
Ding, Z., Granger, C.W.J., Engle, R.E., 1993. Long memory properties of stock market returns and a new model. Journal of Empirical Finance 1, 83–106.
Doornick, J., Ooms, M., 1999. A package for estimating, forecasting and simulating ARFIMA models: ARFIMA package 1.0 for Ox, Mimeo, Nuffield College, Oxford.
Dueker, M., Startz, R., 1998. Maximum likelihood estimation of fractional cointegration with an application to US and Canadian bond rates. Journal of the American Statistical Association 93, 420–426.

Duffie, D., and Singleton, K. (1993), “Simulated Moments Estimation of Markov Models of Asset Prices,” *Econometric*a, 61, 929-952.

Engle, R.F., Granger, C.W.J., 1987. Co-integration and error correction: representation, estimation and testing. Econometrica 55, 251–276.

Engle, R.F., and Lee, G.G.J. (1999), “A Permanent and Transitory Model of Stock Return Volatility,” in R. Engle and H. White (eds.), *Cointegration, Causality, and Forecasting: A Festschrift in Honor of Clive W.J. Grange*r, 475-497. Oxford: Oxford University Press.

Engle, R.F., and Ng, V.K. (1993), “Measuring and Testing the Impact of News on Volatility,” *Journal of Financ*e, 48, 1749-1801.

Fama, E, and French, K, 1988, “Permanent and Temporary Components of Stock Prices”, Journal of Political Economy, 96, 246-273.

Feller, W. (1951), “The Asymptotic Distribution of the Range of Sums of Independent Random Variables,” *Annals of Mathematical Statistic*s, 22, 427-432.

Galant, A.R., Hsieh, D., and Tauchen, G.E. (1997), “Estimation of Stochastic Volatility Models with Diagnostics,” *Journal of Econometric*s, 81, 159-192.

Gallant, A.R., Hsu, C.T., and Tauchen, G.E. (1999), “Calibrating Volatility Diffusions and Extracting Integrated Volatility,” *Review of Economics and Statistic*s, forthcoming.

Garman, M.B., and Klass, M.J. (1980), “On the Estimation of Price Volatility from Historical Data,” *Journal of Busines*s, 53, 67-78.

Geweke, J., Porter-Hudak, S., 1983, The Estimation and Application of Long Memory Time Series Models, Journal of Time Series Analysis, 4, 15-39.

Ghysels, E., Harvey, A., and Renault, E. (1996), “Stochastic Volatility,” in G.S.

Hamilton, J.D. (1994), *Time Series Analysi*s. Princeton: Princeton University Press.

Granger, C.W.J., Ding, Z., 1995. Some properties of absolute return: an alternative measure of risk Annales D'Economie et de Statistique 40, 67–91.

Granger, C.W.J., Joyeux, R., 1980. An introduction to long memory time series models and fractional differencing. Journal of Time Series Analysis 1, 15–29.

Granger, C.W.J., Hyung, N., 1999, Occasional Structural Breaks and Long Memory, University of California, Discussion Paper

Greene, M., and Fielitz, B., 1997, Long-Term Dependence in Common Stock Returns, Journal of Financial Economics, 4, 339-349

Harvey, A., Ruiz, E., and Shephard, N. (1994), “Multivariate Stochastic Variance Models,” *Review of Economic Studie*s, 61, 247-264.

Harvey, A., Ruiz, E., and Shephard, N. (1996), "Estimation of an Asymmetric Stochastic Volatility Model for Asset Returns," *Journal of Business and Economic Statistic*s, 14, 429-434.

Hosking, J.R.M., 1981. Fractional differencing. Biometrika 68, 165–176.

Hull, J., and White, A. (1987), "The Pricing of Options on Assets with Stochastic Volatilities," *Journal of Financ*e, 42, 381-400.

Jegadeesh, N, 1990, "Evidence of Predictable Behavior of Security Returns", Journal of Finance, 45, 881-898.

Johansen, S., 1988. Statistical analysis of cointegration vectors. Journal of Economic Dynamics and Control 12, 231–254.

Kim, S., Shephard, N., and Chib, S. (1998), "Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models," *Review of Economic Studie*s, 65, 361-393.

Lamoureux, C., Lastrapes, W., 1990. Persistence in variance, structural change, and the GARCH model. Journal of Business and Economic Statistics 8, 225–235.

Nelson, D.B. (1991), "Conditional Heteroskedasticity in Asset Pricing: A New Approach," *Econometric*a, 59, 347-370.

Lo, A, 1991, "Long-Term Memory in Stock Market Prices", Econometrica, 59, 1279-1313.

Lo, A and MacKinlay, 1988, "Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test", Review of financial Studies, 1, 41-66.

Mandelbrot, B, 1971, "When Can Price Be Arbitraged Efficiently? A Limit to the Validity of the Random Walk and Martingale Models", The Review of Economics & Statistics, 53, 225-236.

Mandelbrot, B, and Wallis, J, (1968), "Computer Experiments with Fractional Gaussian Noises. Parts 1,2, 3", Water Resources Research, 5, 228-267.

Ooms, M., Doornik, J., 1998. Estimation, simulation and forecasting for fractional autoregressive integrated moving average models, Discussion Paper, Econometric Institute, Erasmus University Rotterdam.

Parkinson, M. (1980), "The Extreme Value Method for Estimating the Variance of the Rate of Return," *Journal of Busines*s, 53, 61-65.

Peng, C.K., S.V. Buldyrev, S. Havlin, M. Simons, H.E. Stanley, A.L. Goldberger, Phys. Rev. E 49 (1994) 1684.

Porteba, J, and Summers, L, 1988, "Mean Reversion in Stock Returns Evidence and Implications", Journal of Financial Economics, 22, 27-60.

Robinson, P.M., 1995a. Log-periodogram regression of time series, with long range dependence Annals of Statistics 23, 1048–1072.
Robinson, P.M., 1995b. Guassian semiparametric estimation of long range dependence. Annals of Statistics 23, 1630–1661.
Rogers, L.C.G., and Satchell, S.E. (1991), “Estimating Variance from High, Low and Closing Prices,” *Annals of Applied Probabilit*y, 1, 504-512.
Ruiz, E. (1994), “Quasi-Maximum Likelihood Estimation of Stochastic Variance Models,” *Journal of Econometric*s, 63, 289-306.
Sandmann, G., and Koopman, S. (1998), “Estimation of Stochastic Volatility Models via Monte Carlo Maximum Likelihood,” *Journal of Econometric*s, 87, 271-301.
Schwert, G.W. (1989), “Why Does Stock Market Volatility Change Over Time,” *Journal of Financ*e, 44, 1115-1153.
Singleton, K. (1997), “Estimation of Affine Asset Pricing Models Using the Empirical Characteristic Function,” Manuscript, Graduate School of Business, Stanford University.
Sentana, E., 1991, Quadratic ARCH Models: A Potential Reinterpretation of ARCH Models as Second-Order Taylor Approximations, Working Paper, London School of Economics and Political Science, London.
Sowell, F., 1989. Maximum likelihood estimation of fractionally integrated time series models, Mimeo, Carnegie-Mellon University.
Sowell, F., 1992. Maximum likelihood estimation of stationary univariate fractionally integrated time series models. Journal of Econometrics 53, 165–188.
Teyssiere, G., 1996. Double long-memory financial time series, Paper No 348, Department of Economics, QMW, London.
Viswanathan, G.M. , S.V. Buldyrev, S. Havlin, H.E. Stanley, Biophys. J. 72 (1997) 866.