PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Markov processes and the distribution of volatility: a comparison of discrete and continuous specifications

S. J. Taylor

Phil. Trans. R. Soc. Lond. A 1999 **357**, 2059-2070
doi: 10.1098/rsta.1999.0417

THE ROYAL SOCIETY

Markov processes and the distribution of volatility: a comparison of discrete and continuous specifications

By S. J. Taylor

Department of Accounting and Finance, University of Lancaster, Lancaster LA1 4YX, UK (S.Taylor@lancaster.ac.uk)

Two mixtures of normal distributions, created by persistent changes in volatility, are compared as models for asset returns. A Markov chain with two states for volatility is contrasted with an autoregressive Gaussian process for the logarithm of volatility. The conditional variances of asset returns are shown to have a bimodal distribution for the former process when volatility is persistent that contrasts with a unimodal distribution for the latter process. A test procedure based upon this contrast shows that a lognormal distribution for sterling/dollar volatility is far more credible than only two volatility states.

Keywords: conditional state probabilities; foreign exchange volatility; Leptokurtic return distributions; Markov chain; mixture distributions; stochastic volatility

1. Introduction

The distributions of asset returns are known to have high peaks, fat tails and excess kurtosis compared with normal distributions. The standard explanation of this empirical phenomenon is that the distribution of returns is a mixture of normal distributions that have different variances. Consider the following factorization of a return r in excess of its mean μ:

$$r - \mu = \sigma u,$$

with $u \sim N(0,1)$ a standardized normal variable and $\sigma > 0$ representing volatility. This factorization is applicable to all the martingale-difference processes for returns that are popular in contemporary research literature, including ARCH models (Bollerslev *et al.* 1994) and stochastic volatility models (Shephard 1996). With the additional assumption that σ and μ are independent, which is common and made in this paper, returns have kurtosis

$$\frac{E[r^4]}{E[r^2]^2} = 3\frac{E[\sigma^4]}{E[\sigma^2]^2} = 3\left(1 + \frac{\text{var}(\sigma^2)}{E[\sigma^2]^2}\right).$$

Returns then have kurtosis in excess of the normal level, 3, whenever the distribution of σ has positive variance.

Many distributions for the volatility variable σ have been proposed and, consequently, many distributions for returns r. Influential examples include a linear function of a Poisson variable for σ^2 (Press 1967), an inverted gamma distribution for

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

σ^2 and a Student distribution for r (Praetz 1972), a lognormal distribution for σ (Clark 1973) and discrete distributions for σ with two possible outcomes (Ball & Torous 1983) or more (Kon 1984). However, there has been relatively little research into methods for inferring the distribution of σ from observed returns.

Two distributions for σ are discussed in this paper and methods are presented that can be used to decide which of the two distributions best describes observed returns. These methods can be extended to make comparisons between more complicated distributions. Section 2 defines the two distributions investigated here: σ has either two possible states or a lognormal distribution. It is shown that these very different distributions for σ can produce very similar distributions for returns. Consequently, progress can only be made by considering stochastic processes for volatility. Section 3 discusses Markov processes for the two-state volatility specification, following Hamilton (1988), and AR(1) processes for the lognormal specification, following Taylor (1986). The corresponding processes for returns are respectively called the 2N-Markov and LNN-AR(1) processes and there are natural extensions to continuous-time processes for prices.

Volatility is a highly persistent process for daily and more frequent returns. Such persistence implies that the probability of the most probable volatility state for the 2N-Markov model, conditional on returns at other times, is often very close to one. A test procedure developed from this result is described in § 4 and investigated for time-series of 2500 simulated daily returns. The test criterion identifies the correct distribution for σ for almost all series when the model parameters have values motivated by empirical studies. Section 5 evaluates the test criterion for 10 years of daily foreign exchange returns for the sterling/dollar rate. The results show that a lognormal distribution for sterling/dollar volatility is far more credible than only two volatility states.

Throughout this paper f is used to represent the probability density function (PDF) of a random variable, which may be continuous or discrete, and ϕ is the density of the standard normal distribution,

$$\phi(x) = (2\pi)^{-1/2} \exp(-\tfrac{1}{2}x^2).$$

2. Univariate mixture distributions

(a) *Two volatility states*

It may be supposed that the volatility σ equals either a lower level $\sigma_{\rm L}$ or a higher level $\sigma_{\rm H}$ with $\sigma_{\rm L} < \sigma_{\rm H}$. Let p be the probability of the lower variance state. Then the PDF of returns is determined by the four parameters μ, $\sigma_{\rm L}$, $\sigma_{\rm H}$, p, and equals

$$f(r) = \frac{p}{\sigma_{\rm L}}\phi\left(\frac{r-\mu}{\sigma_{\rm L}}\right) + \frac{1-p}{\sigma_{\rm H}}\phi\left(\frac{r-\mu}{\sigma_{\rm H}}\right).$$

The moments of this distribution are obtained from $E[\sigma^n] = p\sigma_{\rm L}^n + (1-p)\sigma_{\rm H}^n$ and the assumed independence of volatility and standardized returns. In particular, the returns distribution has mean μ and variance $p\sigma_{\rm L}^2 + (1-p)\sigma_{\rm H}^2$. The kurtosis of both volatility and its logarithm equal $p^{-1}(1-p)^{-1} - 3$, which equals the normal kurtosis of 3 when p is $0.5 \pm \frac{1}{6}\sqrt{3}$, or approximately either 0.211 or 0.789.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

(*b*) *Lognormal volatility*

Alternatively, it may be supposed that the logarithm of volatility is normal with mean α and variance β^2. The PDF of returns is determined by the three parameters μ, α, β, and can only obtained by numerical integration, for example from

$$\begin{aligned} f(r) &= \int f(\sigma) f(r|\sigma)\,\mathrm{d}\sigma \\ &= \int_0^\infty \frac{1}{\sigma^2\beta}\phi\left(\frac{\ln(\sigma)-\alpha}{\beta}\right)\phi\left(\frac{r-\mu}{\sigma}\right)\mathrm{d}\sigma. \end{aligned}$$

The moments of this distribution follow from $E[\sigma^n] = \exp(n\alpha + \frac{1}{2}n\beta^2)$. This returns distribution has mean, variance and kurtosis, respectively, equal to μ, $\exp(2\alpha + 2\beta^2)$ and $3\exp(4\beta^2)$.

(*c*) *Density comparisons and representative parameters*

The PDF of returns is symmetric and fat-tailed for both volatility models. The two models can produce very similar densities for the return r despite the absence of any similarity in the densities for the volatility σ. A typical value for β is 0.4 when volatility is assumed to be lognormal (Taylor 1986, 1994). The unconditional mean and variance of returns can be general for much of this paper and they are, respectively, set to zero and one when comparing models; the representative value for α is then -0.16.

The two-normal (2N) model has one more parameter than the lognormal (LNN) model. Consequently, there are many possible ways to select representative parameters for the former model that produce a close density approximation to the latter model. Representative values for the probability p and the volatility levels σ_{L}, σ_{H} can be defined by equating three volatility moments for the 2N and LNN models; equating values for $E[\sigma^n]$ has the same effect as equating values for $E[|r-\mu|^n]$.

The parameter values used in later sections are obtained by requiring the moments of the two models to be the same for $n = 1, 2, 3$. A simple numerical method provides the solution $p = 0.776$, $\sigma_{\mathrm{L}} = 0.7165$, $\sigma_{\mathrm{H}} = 1.6388$. The kurtoses for the representative parameters are 5.46 for the 2N model and 5.69 for the LNN model.

Figure 1 shows the density functions for the returns distributions defined by the representative parameter values and, for comparison, the PDF of the standard normal distribution. The maximum difference between the density functions of returns for the 2N and LNN models is 0.0206 and occurs at the modes. The peak of the LNN density is slightly higher than the peak of the 2N density and both these peaks are well above the peak of the normal (N) density. Figure 2 shows differences between cumulative distribution functions. The maximum difference for the 2N and LNN distributions is only 0.0052, at $r = \pm 1.23$. This is much more than the maximum difference for the LNN and N distributions, that is 0.0404 at $r = \pm 0.63$.

Equating values for $E[\sigma^4]$ instead of $E[\sigma^3]$ provides similar parameters and only a slightly worse match between the densities of returns. A better match is obtained by using $E[\ln\sigma]$, but p is then 0.639 and this value is less than typical values discussed in previous literature. The ratio $\sigma_{\mathrm{H}}/\sigma_{\mathrm{L}}$ is between 2.2 and 2.4 for all two-normal parameters that have been obtained by matching the moments of the lognormal model when β is 0.4.

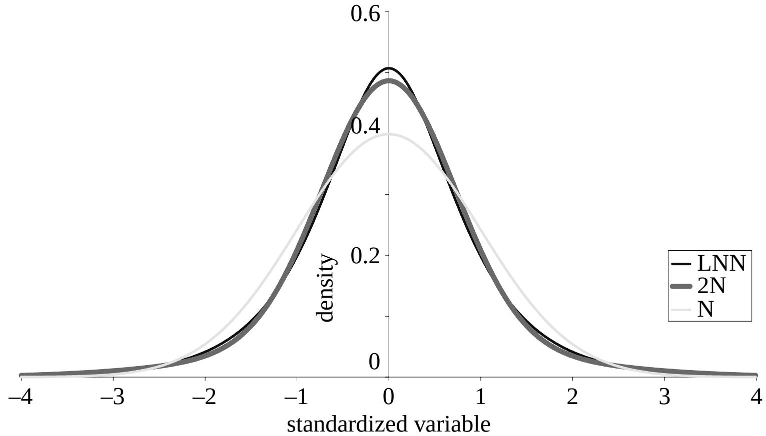

Figure 1. Density functions.

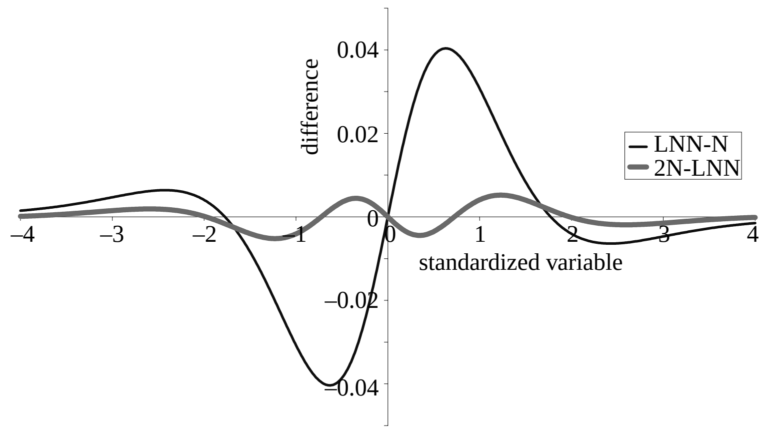

Figure 2. Differences between cumulative distribution functions.

The similarity between the fat-tailed distributions for returns exhibited in figure 1 shows that it will be difficult to use properties of univariate return distributions to decide which of the two volatility models provides the best description of observed returns. The task appears to be futile for samples of 2500 daily returns, and consequently it is necessary to consider multivariate distributions.

3. Multivariate mixture distributions

Next consider specifications of stationary stochastic processes for n consecutive returns $\{r_t, 1 \leqslant t \leqslant n\}$. Without loss of generality, expected returns are assumed to be zero. It is well known that appropriate processes have positive autocorrelations for functions of $|r_t|$ and that this occurs when there is positive dependence in the process for volatility (Taylor 1986).

Let $\rho_{s,\tau} = \operatorname{corr}(s_t, s_{t+\tau})$, $\tau > 0$, denote the autocorrelations of a stationary process $\{s_t\}$. The autocorrelations of $|r_t| = \sigma_t|u_t|$ are proportional to the autocorrelations of volatility σ_t as follows (Taylor 1986, p. 73):

$$\rho_{|r|,\tau} = \frac{2(A-1)}{\pi A - 2}\rho_{\sigma,\tau} \quad \text{with } A = \frac{E[\sigma^2]}{E[\sigma]^2}.$$

(*a*) *Lognormal volatility*

The process for $\ln(\sigma_t)$ that has received the most attention is the Gaussian first-order autoregressive process, with mean α, variance β^2 and autocorrelations Φ^τ. Then

$$A = \exp(\beta^2),$$
$$\rho_{\sigma,\tau} = \frac{[\exp(\beta^2\Phi^\tau) - 1]}{[\exp(\beta^2) - 1]} \cong \Phi^\tau.$$

Representative values for β and Φ are 0.4 and 0.98, respectively. The first autocorrelations for the processes $|r_t|$, $\ln(\sigma_t)$ and σ_t are then 0.201, 0.98 and 0.978, respectively. The derived model for returns is called the LNN-AR(1) model in this paper. The mathematical properties of the returns model were first discussed in Taylor (1982) and developed further in Taylor (1986). See Shephard (1996) for a recent survey of the LNN-AR(1) model and extensions.

(*b*) *Two volatility states*

The logical specification of a stochastic process for two-state volatility random variables is a Markov chain for $\{\sigma_t\}$ with transition probabilities

$$q_{\mathrm{LH}} = P(\sigma_{t+1} = \sigma_{\mathrm{H}} \mid \sigma_t = \sigma_{\mathrm{L}}),$$
$$q_{\mathrm{HL}} = P(\sigma_{t+1} = \sigma_{\mathrm{L}} \mid \sigma_t = \sigma_{\mathrm{H}}).$$

The probability of the lower volatility state, p, and the transition probabilities are constrained by the formula $pq_{\mathrm{LH}} = (1-p)q_{\mathrm{HL}}$. The autocorrelations of the volatility process are given by

$$\rho_{\sigma,\tau} = \psi^\tau,$$
$$\psi = 1 - (q_{\mathrm{LH}} + q_{\mathrm{HL}}) = 1 - \frac{q_{\mathrm{LH}}}{1-p} = 1 - \frac{q_{\mathrm{HL}}}{p}.$$

Representative values for p and ψ are 0.776 and 0.98, respectively, and then q_{LH} and q_{HL} are 0.004 48 and 0.015 52, respectively. Changes of state are infrequent because the volatility process is highly persistent. In the lower volatility state, the expected time until the next change of state is $1/q_{\mathrm{LH}} = 223$ time units. By also selecting $\sigma_{\mathrm{L}} = 0.7165$ and $\sigma_{\mathrm{H}} = 1.6388$, the autocorrelations of $|r_t|$ are almost the same as for the autoregressive lognormal model.

The model for returns derived from the Markov chain for volatility is here called the 2N-Markov model. Hamilton (1988) provides the first analysis of this model, and its forecasting potential is studied in Pagan & Schwert (1990). Theoretical and empirical results for the general Markov model having two or more volatility states are provided by Ryden *et al.* (1998).

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

(c) Conditional probabilities for two states

The probabilities of the volatility states conditional upon sets of returns are used extensively in this paper. Let $B_t = \{r_T, 1 \leqslant T < t\}$ be all returns observed *before* time t. In the following equations, the summations $\sum$ are over the two possible outcomes for σ_{t-1}. The conditional probability distribution of σ_t given B_t is

$$\begin{aligned} f(\sigma_t \mid B_t) &= \sum \frac{f(\sigma_{t-1}, \sigma_t, B_t)}{f(B_t)} \\ &= \sum \frac{f(B_{t-1}) f(\sigma_{t-1} \mid B_{t-1}) f(r_{t-1} \mid \sigma_{t-1}, B_{t-1}) f(\sigma_t \mid \sigma_{t-1}, B_t)}{f(B_t)} \\ &= \frac{\sum f(\sigma_{t-1} \mid B_{t-1}) f(r_{t-1} \mid \sigma_{t-1}) f(\sigma_t \mid \sigma_{t-1})}{\sum f(\sigma_{t-1} \mid B_{t-1}) f(r_{t-1} \mid \sigma_{t-1})}. \end{aligned}$$

Consequently, the conditional probabilities of the lower state given past returns, denoted

$$p_t^B = P(\sigma_t = \sigma_{\mathrm{L}} \mid B_t),$$

can be calculated recursively from p_{t-1}^B, r_{t-1} and the parameters σ_{L}, σ_{H}, q_{LH} and q_{HL} that define the volatility process, commencing with $p_1^B = p$.

Note that the likelihood of a set of returns can be calculated without difficulty from the product of conditional densities

$$f(r_t \mid B_t) = \sum_{\sigma_t = \sigma_{\mathrm{L}}, \sigma_{\mathrm{H}}} f(r_t \mid \sigma_t) f(\sigma_t \mid B_t),$$

that are functions of the probabilities p_t^B, the returns r_t and the volatility parameters.

Let $A_t = \{r_T, t < T \leqslant n\}$ be all returns observed *after* time t. As all the processes are reversible, it is straightforward to derive the further probabilities

$$p_t^A = P(\sigma_t = \sigma_{\mathrm{L}} \mid A_t)$$

recursively from p_{t+1}^A, r_{t+1} and the volatility parameters. The conditional probabilities given all returns except the return at time t follow from

$$\begin{aligned} f(\sigma_t \mid A_t, B_t) &= \frac{f(B_t) f(\sigma_t \mid B_t) f(A_t \mid \sigma_t)}{f(A_t, B_t)} \\ &= \frac{f(B_t) f(\sigma_t \mid B_t) f(A_t) f(\sigma_t \mid A_t)}{f(A_t, B_t) f(\sigma_t)}. \end{aligned}$$

Hence it can be shown that

$$p_t^{AB} = P(\sigma_t = \sigma_{\mathrm{L}} \mid A_t, B_t) = \frac{(1-p) p_t^A p_t^B}{(1-p) p_t^A p_t^B + p(1-p_t^A)(1-p_t^B)}.$$

(d) Examples of conditional probabilities

Time-series of 2500 returns have been simulated from both the 2N-Markov and the LNN-AR(1) models, using the representative parameters, followed by calculation of the conditional probabilities p_t^B, p_t^A and p_t^{AB}. These probabilities are often very near either 0 or 1 because of the high persistence in the simulated processes; the

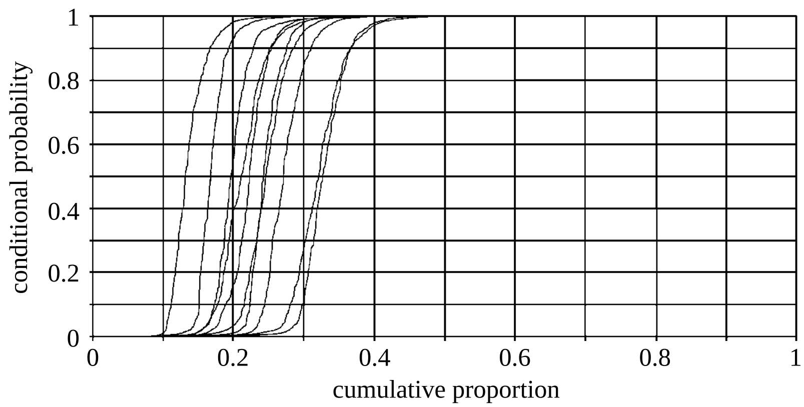

Figure 3. Ranked probabilities for the 2N-Markov model.

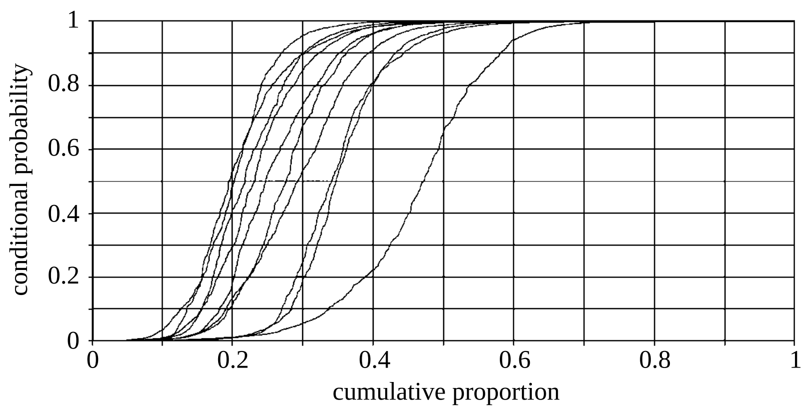

Figure 4. Ranked probabilities for the LNN-AR(1) model.

extreme values of p_t^{AB} are 0.000 07 and 0.999 78. The probabilities p_t^{AB} are more often near 0 or 1 than the probabilities p_t^B. Only 14.7% of the p_t^{AB} are in the interval $[0.01, 0.99]$ for the 2N-Markov model, with 16.6% below 0.01 and 68.7% above 0.99. The corresponding frequencies for the LNN-AR(1) model are 36.2% in the interval $[0.01, 0.99]$, 13.1% below 0.01 and 50.7% above 0.99.

Figure 3 illustrates the cumulative distributions of the probabilities p_t^{AB} for 10 simulations of 2500 returns for the 2N-Markov model. The points on the figure show the proportions of the p_t^{AB}, measured on the horizontal axis, that are less than levels measured on the vertical axis; p_t^{AB} is measured on the vertical axis and its cumulative distribution function on the horizontal axis. It can be seen that the proportions of very low and very high p_t^{AB} vary substantially from series to series. This is a consequence of the high volatility persistence, which causes the proportions of time spent in the two volatility states to vary substantially across the 10 series.

Figure 4 shows the same information for 10 simulations of the LNN-AR(1) model. The slopes of the central sections of the curves are much less steep on figure 4 than on figure 3, reflecting the higher occurrence of probabilities p_t^{AB} between 0.01 and 0.99.

4. Criteria that distinguish the volatility models

(a) The problem

Now consider the problem of identifying the volatility model, given a time-series of observed returns and the assumption that the returns process is either 2N-Markov or LNN-AR(1). This problem has been investigated for series of 2500 returns with the objective of making the correct identification for 99% of simulated series. It is tempting to suppose that the high frequency of probabilities p_t^{AB} near to 0 or 1 for the 2N-Markov model will enable the problem to have a simple solution. Plausible test statistics can be constructed from (a) the kurtosis of those returns for which $p_t^{AB} > 0.999$; (b) the first-lag autocorrelation of the sub-series $\{|r_t|\}$ for which $p_t^{AB} > 0.999$; and (c) the proportion of the p_t^{AB} in intervals such as $[0.01, 0.99]$. However, these and other simple methods do not identify the correct volatility model with sufficient accuracy, particularly when it is necessary to estimate the model parameters from the data.

A likelihood-ratio can be computed for the two volatility models after maximizing two likelihood functions. There are no difficulties for the 2N-Markov model for which an EM algorithm can be used (Hamilton 1990). However, sophisticated numerical algorithms are required for the LNN-AR(1) model (see Danielsson 1994; Shephard 1996; Kim *et al*. 1998). The likelihood functions are not nested and they have different numbers of parameters. It is quite probable that most people will prefer the simpler approach that is now described.

(b) The conditional variance test

Let h_t denote the variance of a return conditional upon previous returns. For the 2N-Markov model these variances are simply

$$h_t^{(2\mathrm{N})} = p_t^B \sigma_\mathrm{L}^2 + (1 - p_t^B)\sigma_\mathrm{H}^2,$$

and they are constrained to be in the interval $[\sigma_\mathrm{L}^2, \sigma_\mathrm{H}^2]$. The distribution of the variances is bimodal, with most outcomes close to either σ_L^2 or σ_H^2. In contrast, the conditional variances for the LNN-AR(1) model are unimodal and only constrained to be positive. Consequently, the two models will provide clearly different conditional variances from the same data. As the conditional variances from the wrong model contain no incremental information beyond that supplied by the correct model, an encompassing regression test should be able to identify the correct model.

There are three initial steps in the proposed test. First, data $\{r_t\}$ are standardized to have zero mean and unit variance. Second, the four parameters of the 2N-Markov model, σ_L, σ_H, p, ψ, are estimated by maximizing the likelihood of the standardized data, with the unconditional variance constrained to be unity, so $p\sigma_\mathrm{L}^2 + (1-p)\sigma_\mathrm{H}^2 = 1$. Third, the three parameters of the GARCH(1,1) model, known to have very similar mathematical properties to the LNN-AR(1) model (Taylor 1994), are estimated by quasi-maximum likelihood after supposing the conditional distributions of returns are $N(0, h_t)$ with $h_t = c + ar_{t-1}^2 + bh_{t-1}$; the non-negative parameters a, b, c are constrained to have $a + b + c = 1$.

The parameter estimates are then used to obtain conditional variances for the standardized returns r_t. These are conditioned either on previous returns B_t or on

Table 1.

model	identifications		average estimates of	
	correct	wrong	β_1	β_2
2N-Markov	2497	3	1.05	−0.05
LNN-AR(1)	2308	192	0.30	0.70

previous and future returns $B_t + A_t$. In the latter case, for the GARCH specification, the conditional variance is approximated by the average of $\text{var}(r_t \mid B_t)$ and $\text{var}(r_t \mid A_t)$; the terms $\text{var}(r_t \mid A_t)$ are obtained from the approximation

$$h_t = c + ar_{t+1}^2 + bh_{t+1}.$$

Given time-series of conditional variances $h_t^{(2\text{N})}$ and $h_t^{(\text{GARCH})}$, the final step is to estimate the regression

$$E[r_t^2] = \beta_0 + \beta_1 h_t^{(2\text{N})} + \beta_2 h_t^{(\text{GARCH})}$$

by ordinary least squares. The volatility model is then identified as 2N-Markov if and only if $\beta_1 > \beta_2$. There are doubtless many ways to enhance the test procedure. Here the emphasis is on providing a straightforward method that avoids computational complexity.

(*c*) *Performance of the test*

The test has been evaluated for 5000 simulated series of 2500 returns. Half of the series are generated by the 2N-Markov model and the other half by the LNN-AR(1) model. The representative parameters are used to obtain the simulated series, that are then standardized and used to obtain estimated parameters and conditional variances. The first 50 returns are omitted from the regression when variances are conditioned on B_t and the last 50 are also omitted when conditioning upon $B_t + A_t$.

The information in table 1 summarizes the results when the variances are conditioned on B_t. In very few cases, a mistake is made about data generated by the 2N-Markov model. The error rate for the LNN-AR(1) model is not satisfactory and is caused by the high average weight given to β_1.

Satisfactory results are obtained when the variances are conditioned on $B_t + A_t$, as shown in table 2. The overall error rate is then $64/5000 = 1.28\%$. Few series identify the model incorrectly for both the test using information B_t and the test using information $B_t + A_t$. The two tests give the same result for 4775 of the 5000 series and they give conflicting results for the other 225 series. Only 17 of the 4775 series that give the same identification provide a wrong identification. The error rate, conditional on the same identification, is less than 0.4% for the simulation experiment.

Figures 5 and 6 shows the estimates of (β_1, β_2), obtained using the information $B_t + A_t$, respectively, for the 2N-Markov and LNN-AR(1) models. The estimates are scattered across large regions, particularly for the 2N-Markov model, although 98% of the estimates are on the correct side of the line $\beta_1 = \beta_2$ on figure 5 and more than 99% on figure 6.

Table 2.

model	identifications		average estimates of	
	correct	wrong	β_1	β_2
2N-Markov	2451	49	1.10	−0.13
LNN-AR(1)	2485	15	0.15	0.97

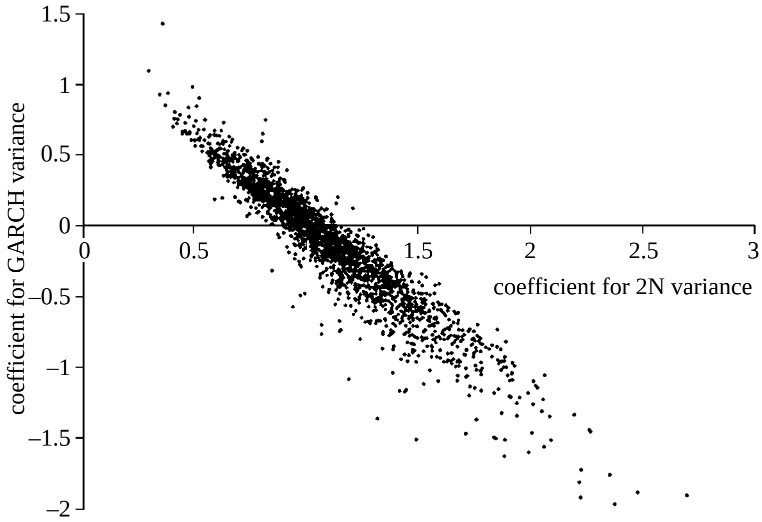

Figure 5. Estimated coefficients for the 2N-Markov model.

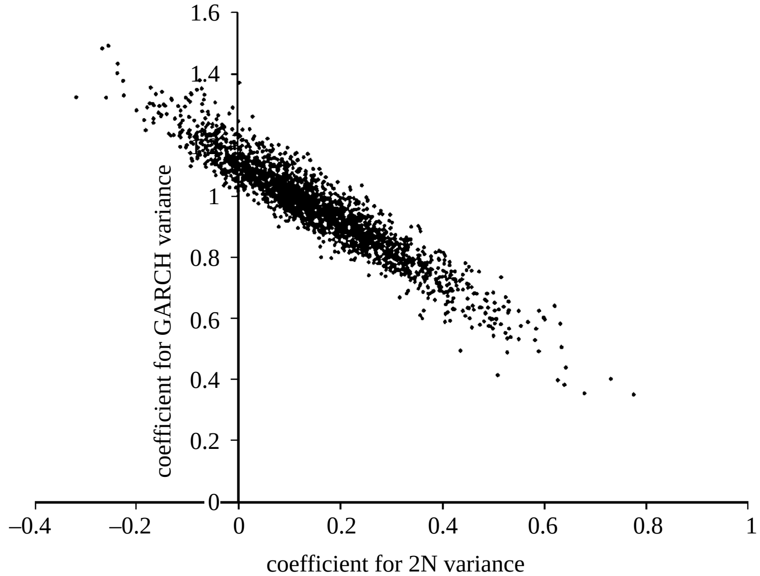

Figure 6. Estimated coefficients for the LNN-AR(1) model.

5. Results for the sterling/dollar rate

The conditional variance test has been performed for 2529 sterling/dollar returns from December 1982 to November 1991 inclusive, calculated from Chicago futures prices p_t as $r_t = \ln(p_t/p_{t-1})$. The estimates of (β_1, β_2) are $(0.19, 0.90)$ using previous returns to calculate the conditional variances, and $(0.04, 1.28)$ using previous and future returns. As $\beta_2 > \beta_1$ for both regressions, these results favour the LNN-AR(1) model.

Foreign exchange returns are known, however, to be more variable following weekends; Chicago futures returns are also more variable following US holidays (Taylor 1994). To control for these effects, a second set of estimates are obtained from returns that are standardized according to six categories: returns that include a holiday period, Monday returns that do not include a holiday, etc. The estimates of (β_1, β_2) change by small amounts, to $(0.15, 0.94)$ and $(0.02, 1.30)$, for conditional variances obtained first from previous returns and second from previous and future returns. The estimates $(0.02, 1.30)$ are to the left of all the dots on figure 5 for the 2N-Markov model, while they are not far from the central group of dots on figure 6 for the LNN-AR(1) model, although $\beta_1 + \beta_2 = 1.32$ is relatively high. The empirical evidence is overwhelmingly in favour of the LNN-AR(1) model if a choice must be made between this model and the 2N-Markov model.

The parameter estimates for the 2N-Markov model, estimated from returns standardized by day of the week, are $\sigma_{\mathrm{L}} = 0.756$, $\sigma_{\mathrm{H}} = 1.461$, $p = 0.726$ and $\psi = 0.880$. The ratio $\sigma_{\mathrm{H}}/\sigma_{\mathrm{L}}$ is estimated as 1.93, less than the 2.29 for the representative parameters, while the estimate of p is only slightly less than the representative level of 0.776. The kurtosis is only 4.46 for the estimated model, compared with 5.46 for the representative parameters and 5.87 for the empirical data. The estimate of the persistence parameter ψ is surprisingly low and less than the estimates obtained from simulation of the 2N-Markov and LNN-AR(1) models. However, the maximum log-likelihood as a function of ψ is similar for a wide range of values for ψ; a two-sided 95% confidence interval includes 0.95.

The estimates of the GARCH(1,1) parameters are $a = 0.042$ and $b = 0.943$, giving a persistence estimate of 0.985, which is similar to the representative level of 0.98. For the LNN distribution, matching the moments of absolute standardized returns gives an estimate of β equal to 0.411, which is close to the representative level of 0.4. The fitted kurtosis is then 5.90 and thus very near to the empirical quantity, 5.87.

6. Concluding remarks

To make inferences about the distribution of volatility it can be necessary and sufficient to consider properties of multivariate return distributions. The high positive correlation between volatility on nearby days, known as volatility persistence, makes it possible to calculate conditional variances whose distribution provides useful information about the distribution of volatility. In particular, reliable choices can be made between a two-state distribution and a lognormal distribution for volatility. More complicated distributions for volatility may be more appropriate. One possibility is that returns are a mixture of three (or more) normal distributions. Another is that returns are a mixture of two (or more) Student-t distributions whose scale parameters are determined either by a discrete or a lognormal distribution. All of these and

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

many other possibilities can be investigated by adapting the methods described in this paper.

References

Ball, C. A. & Torous, W. N. 1983 A simplified jump process for common stock returns. *J. Financial Quantitative Analysis* **18**, 53–65.

Bollerslev, T., Engle, R. F. & Nelson, D. B. 1994 ARCH models. In *Handbook of econometrics*, vol. IV, pp. 2959–3038. New York: North-Holland.

Clark, P. K. 1973 A subordinated stochastic process model with finite variance for speculative prices. *Econometrica* **41**, 135–155.

Danielsson, J. 1994 Stochastic volatility in asset prices: estimation with simulated maximum likelihood. *J. Econometrics* **61**, 375–400.

Hamilton, J. D. 1988 Rational-expectations econometric analysis of changes in regime: an investigation of the term structure of interest rates. *J. Economic Dynamics Control* **12**, 385–423.

Hamilton, J. D. 1990 Analysis of time series subject to changes in regime. *J. Econometrics* **45**, 39–70.

Kim, S., Shephard, N. & Chib, S. 1998 Stochastic volatility: likelihood inference and comparison with ARCH models. *Rev. Economic Studies* **65**, 361–393.

Kon, S. J. 1984 Models of stock returns—a comparison. *J. Finance* **39**, 147–165.

Pagan, A. R. & Schwert, G. W. 1990 Alternative models for conditional stock volatility. *J. Econometrics* **45**, 267–290.

Praetz, P. D. 1972 The distribution of share price changes. *J. Business* **45**, 49–55.

Press, S. J. 1967 A compound events model for security prices. *J. Business* **40**, 317–335.

Ryden, T., Terasvirta, T. & Asbrink, S. 1998 Stylized facts of daily return series and the hidden Markov model. *J. Appl. Econometrics* **13**, 217–244.

Shephard, N. 1996 Statistical aspects of ARCH and stochastic volatility. In *Likelihood, time series with econometric and other applications* (ed. D. R. Cox, D. V. Hinkley & O. E. Barndorff-Nielsen), pp. 1–67. London: Chapman & Hall.

Taylor, S. J. 1982 Financial returns modelled by the product of two stochastic processes, a study of daily sugar prices 1961–79. In *Time series analysis: theory and practice 1* (ed. O. D. Anderson), pp. 203–226. New York: North-Holland.

Taylor, S. J. 1986 *Modelling financial time series*. Chichester: Wiley.

Taylor, S. J. 1994 Modeling stochastic volatility: a review and comparative study. *Math. Finance* **4**, 183–204.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES