Multivariate Stochastic Variance Models
Author(s): Andrew Harvey, Esther Ruiz, Neil Shephard
Source: *The Review of Economic Studies*, Vol. 61, No. 2 (Apr., 1994), pp. 247-264
Published by: Oxford University Press
Stable URL: http://www.jstor.org/stable/2297980
Accessed: 03/05/2011 13:11

Review of Economic Studies (1994) **61**, 247-264
0034-6527/94/00130247$02.00

Multivariate Stochastic Variance Models

ANDREW HARVEY
London School of Economics

ESTHER RUIZ
University Carlos III de Madrid

and

NEIL SHEPHARD
Nuffield College, Oxford

First version received May 1992; *final version accepted November* 1993 (*Eds.*)

Changes in variance, or volatility, over time can be modelled using the approach based on autoregressive conditional heteroscedasticity (ARCH). However, the generalizations to multivariate series can be difficult to estimate and interpret. Another approach is to model variance as an unobserved stochastic process. Although it is not easy to obtain the exact likelihood function for such stochastic variance models, they tie in closely with developments in finance theory and have certain statistical attractions. This article sets up a multivariate model, discusses its statistical treatment and shows how it can be modified to capture common movements in volatility in a very natural way. The model is then fitted to daily observations on exchange rates.

1. INTRODUCTION

Many financial time series, such as stock returns and exchange rates, exhibit changes in volatility over time. These changes tend to be serially correlated and in the *generalized autoregressive conditional heteroscedasticity*, or GARCH model, developed by Engle (1982) and Bollerslev (1986), such effects are captured by letting the conditional variance be a function of the squares of previous observations and past variances. Since the model is formulated in terms of the distribution of the one-step ahead prediction error, maximum likelihood estimation is straightforward. A wide range of GARCH models have now appeared in the econometric literature; see the recent survey by Bollerslev *et al.* (1992).

An alternative approach is to set up a model containing an unobserved variance component, the logarithm of which is modelled directly as a linear stochastic process, such as an autoregression. Models of this kind are called *stochastic volatility models* or *stochastic variance* (*SV*) *models*. They are the natural discrete-time versions of the continuous-time models upon which much of modern finance theory, including generalizations of the Black–Scholes result on option pricing, has been developed; see, for example, Hull and White (1987) and Taylor (1993). Their principal disadvantage is that they are difficult to estimate by maximum likelihood. However they do have other compensating statistical attractions; for example their properties are easily obtained from the properties of the process generating the variance component. Furthermore, they generalize to multivariate series in a very natural way. The main aim of this article is to show how multivariate stochastic variance models can be handled statistically and to explore how well they fit real data.

Section 2 reviews some of the basic ideas of univariate GARCH and SV models and compares their properties. The estimation of SV models by a quasi-maximum likelihood procedure is then discussed. In Section 3 it is shown how multivariate SV models can be formulated, and how they compare with multivariate GARCH models. The way in which they can handle common movements in volatility in different series is described in Section 4 and this is related to ideas of co-integration in variance. Section 5 presents an example in which the model is fitted to four sets of exchange rates, and Section 6 generalizes the methods so as to handle heavy tailed distributions. The conclusions are given in Section 7.

2. UNIVARIATE MODELS

Let the series of interest be made up of a Gaussian white noise process, with unit variance, multiplied by a factor σ_t, the standard deviation, that is

$$y_t=\sigma_t\varepsilon_t, \qquad t=1,\dots,T, \qquad \varepsilon_t\sim \mathrm{NID}(0,1). \tag{1}$$

In the GARCH(1, 1) model,

$$\sigma_t^2=\gamma+\alpha y_{t-1}^2+\beta\sigma_{t-1}^2, \qquad \gamma>0, \qquad \alpha+\beta<1. \tag{2}$$

This may be generalized by adding more lags of both the squared observations and the variance. All GARCH models are martingale differences, and if the sum of the α and β coefficients is less then one, they have constant finite variance and so are white noise. However, obtaining the general conditions under which σ_t^2 is positive and y_t is stationary is not straightforward; see, for example, Nelson and Cao (1992) and Bollerslev and Engle (1993). Similarly, although it can be shown that y_t exhibits excess kurtosis, the necessary restrictions for fourth moments to exist are not easy to derive.

The dynamics of a GARCH model show up in the autocorrelation function (ACF) of the squared observations. In the GARCH(1, 1) case, the ACF is like that of an ARMA(1, 1) process. If the sum of α and β is close to one, the ACF will decay quite slowly, indicating a relatively slowly changing conditional variance. This has often been observed to happen in practice, and GARCH(1, 1) models with $\alpha+\beta$ close to unity are quite common with real data.

If $\alpha+\beta$ is set to one in the GARCH(1, 1) model, it is no longer weakly stationary since it does not have finite variance. However, Δy_t^2 is stationary and has an ACF like that of an MA(1) process, indicating an analogy with the ARIMA(0, 1, 1) process. This model is called *integrated GARCH*, or *IGARCH*; see Engle and Bollerslev (1986). It does not follow, though, that y_t^2 will behave like an integrated process in all respects, and, in fact, Nelson (1990) has shown that σ_t^2 is strictly stationary.

The IGARCH model is still a martingale difference, and so forecasts of all future observations are zero. If γ is positive, the predicted variances increase with the lead time. On the other hand, if γ is set to zero, Nelson (1990) shows that the IGARCH process has the rather strange property that, no matter what the starting point, σ_t^2 tends towards zero for most parameter values, so that the series effectively disappears. This leads him to suggest a model in which $\log\sigma_t^2$ has the characteristics of a random walk; see Nelson (1991). This is an example of an *exponential ARCH*, or *EGARCH*, model. Such models have the additional attraction that they can be shown to be a discrete-time approximation to some of the continuous-time models of finance theory.

In a stochastic variance model for (1), the logarithm of σ_t^2, denoted h_t, is modelled as a stochastic process. As with EGARCH, working in logarithms ensures that σ_t^2 is always positive, but the difference is that it is not directly observable. A simple model for h_t is the AR(1) process

$$h_t = \gamma + \varphi h_{t-1} + \eta_t, \qquad \eta_t \sim \mathrm{NID}(0, \sigma_\eta^2), \tag{3}$$

with η_t generated independently of ε_s for all t, s. Equation (3) is the natural discrete-time approximation to the continuous-time Orstein–Uhlenbeck process used in finance theory. Dassios (1992) has shown that (1) and (3) is a better discrete-time approximation to the model used in Hull and White (1987) than an EGARCH model. More specifically, if δ denotes the distance between observations, he shows that the density of the variance process converges to the density of the continuous-time variance process at rate δ, whereas in the case of EGARCH the convergence is at rate $\delta^{1/2}$. Similar convergence results hold for the joint moments of the observations.

If $|\varphi| < 1$, we know from standard theory that h_t is strictly stationary, with mean $\gamma_h = \gamma/(1-\varphi)$ and variance $\sigma_h^2 = \sigma_\eta^2/(1-\varphi^2)$. Since y_t is the product of two strictly stationary processes, it must also be strictly stationary. Thus the restrictions needed to ensure stationarity of y_t, both in the strict and weak sense, are just the standard restrictions needed to ensure stationarity of the process generating h_t.

The fact that y_t is white noise follows almost immediately when the two disturbance terms are mutually independent. The odd moments of y_t are all zero because ε_t is symmetric. The even moments can be obtained by making use of a standard result for the lognormal distribution, which in the present context tells us that since $\exp(h_t)$ is lognormal, its j-th moment about the origin is $\exp\{j\gamma_h + \frac{1}{2}j^2\sigma_h^2\}$. It follows almost immediately that the variance of y_t is $\exp\{\gamma_h + \frac{1}{2}\sigma_h^2\}$. Similarly the kurtosis is $3\exp\{\sigma_h^2\}$, which is greater than three when σ_h^2 is positive; see Taylor (1986, Chapter 3). Unlike a GARCH model, the fourth moment always exists when h_t is stationary.

The dynamic properties of the model appear most clearly in $\log y_t^2$. Since $y_t = \varepsilon_t \exp(\frac{1}{2}h_t)$,

$$\log y_t^2 = h_t + \log \varepsilon_t^2, \qquad t = 1, \ldots, T. \tag{4}$$

The mean and variance of $\log \varepsilon_t^2$ are known to be $-1{\cdot}27$ and $\pi^2/2 = 4.93$, respectively; see Abramovitz and Stegun (1970, p. 943). Thus $\log y_t^2$ is the sum of an AR(1) component and white noise and so its ACF is equivalent to that of an ARMA(1, 1). Its properties are therefore similar to those of GARCH(1, 1). Indeed, if σ_h^2 is small and/or φ is close to one, y_t^2 behaves approximately as an ARMA(1, 1) process; see Taylor (1986, p. 74–5, 1993).

The model can be generalised so that h_t follows any stationary ARMA process, in which case y_t is also stationary and its properties can be deduced from the properties of h_t. Alternatively h_t can be allowed to follow a random walk

$$h_t = h_{t-1} + \eta_t, \qquad \eta_t \sim \mathrm{NID}(0, \sigma_\eta^2). \tag{5}$$

In this case $\log y_t^2$ is a random walk plus noise, and the best linear predictor of the current value of h_t is an exponentially weighted moving average (EWMA) of past values of $\log y_t^2$. Thus there is a parallel with the IGARCH model where the conditional variance is also an EWMA. The crucial difference is that while the IGARCH conditional variance is known exactly, the variance generated by (5) is an unobserved component, and a better estimator can be obtained by making use of subsequent observations.

The SV model with h_t following a random walk is clearly non-stationary, with $\log y_t^2$ being stationary after differencing. It is quite close to an EGARCH model in this respect. There is no need to introduce a constant term to prevent the kind of behaviour demonstrated for IGARCH by Nelson. As a result the model contains only one unknown parameter.

The estimation of SV models has usually been carried out by variants of the method of moments; see, for example, Scott (1987), Chesney and Scott (1989), Melino and Turnbull (1990) and the references in Taylor (1993). The approach proposed here is a quasi-maximum likelihood method, computed using the Kalman filter. It was put forward independently by Nelson (1988).

In order to estimate the parameters, φ, γ and σ_η^2, consider the following state-space model obtained from (3) and (4):

$$\log y_t^2 = -1\cdot 27 + h_t + \xi_t \tag{6a}$$

$$h_t = \gamma + \varphi h_{t-1} + \eta_t \tag{6b}$$

where, $\xi_t = \log \varepsilon_t^2 + 1\cdot 27$ and $\text{Var}(\xi_t) = \pi^2/2$. The general form of the model allows for the possibility of the original disturbances in (1) and (3) being correlated. Nevertheless in (6), ξ_t and η_t are uncorrelated; see Appendix A. The question of taking account of any correlation between the original disturbances is to be examined in a later paper.

Although the Kalman filter can be applied to (6), it will only yield minimum mean square *linear* estimators (MMSLEs) of the state and future observations rather than MMSEs. Furthermore, since the model is not conditionally Gaussian, the exact likelihood cannot be obtained from the resulting prediction errors. Nevertheless estimates can be obtained by treating ξ_t as though it were NID(0, $\pi^2/2$) and maximizing the resulting quasi-likelihood function. Asymptotic standard errors, which take account of the specific form of the non-normality in ξ_t, can be computed using the results established by Dunsmuir (1979, p. 502). The experiments reported in Ruiz (1994) suggest that his QML method works well for the sample sizes typically encountered in financial economics and is usually to be preferred to the corresponding method of moments estimator. A further attraction of applying QML to SV models is that the assumption of normality for ε_t can be relaxed, in which case σ_ξ^2 is estimated unrestrictedly; see Section 6.

The Kalman filter approach is still valid when φ is one. The only difference is that the first observation is used to initialize the Kalman filter, whereas when $|\varphi| < 1$ the unconditional distribution of h_t is available at $t = 0$. Once the parameters have been estimated, predictions of future volatility can be made from the predictions of $\log y_t^2$. A smoother can be used to estimate volatility within the sample period; this is also done by Melino and Turnbull (1990) and Scott (1987).

3. MULTIVARIATE MODELS

The multivariate GARCH model, set out in Bollerslev, Engle and Wooldridge (1988), can, in principle, be estimated efficiently by maximum likelihood. However, the number of parameters can be very large, so it is usually necessary to impose restrictions. For example, Bollerslev (1990) proposes a representation in which the conditional correlations are assumed to be constant. This assumption considerably simplifies estimation and inference, and, according to the evidence in Baillie and Bollerslev (1990) and Schwert and Seguin (1990), it is often empirically reasonable.

Stochastic variance models generalize to multivariate series as follows. Let y_t be an $N\times 1$ vector, with elements

$$y_{it}=\varepsilon_{it}(\exp\{h_{it}\})^{1/2}, \qquad i=1,\ldots,N, \qquad t=1,\ldots,T, \tag{7}$$

where y_{it} is the observation at time t of series i, and $\varepsilon_t=(\varepsilon_{1t},\ldots,\varepsilon_{Nt})'$ is a multivariate normal vector with zero mean and a covariance matrix, Σ_ε, in which the elements on the leading diagonal are unity and the off-diagonal elements are denoted as ρ_{ij}. Following (3), the variances may be generated by AR(1) processes

$$h_{it}=\gamma_i+\varphi_i h_{it-1}+\eta_{it}, \qquad i=1,\ldots,N, \tag{8}$$

where $\eta_t=(\eta_{1t},\ldots,\eta_{Nt})'$ is multivariate normal with zero mean and covariance matrix Σ_η. The model in (7) does not allow the covariances to evolve over time independently of the variances. Thus it is restricted in a similar way to the constant conditional correlation GARCH model.[1]

Model (8) could be generalized so that the $N\times 1$ vector h_t is a multivariate AR(1) process or even an ARMA process. Although the properties of such models could be derived relatively easily, generalizations of this kind are probably not necessary in practice. We will instead focus attention on the special case when h_t is a multivariate random walk. Transforming as in (6) gives

$$w_t=-1{\cdot}27\iota+h_t+\xi_t \tag{9a}$$

$$h_t=h_{t-1}+\eta_t \tag{9b}$$

where w_t and ξ_t are $N\times 1$ vectors with elements $w_{it}=\log y_{it}^2$ and $\xi_{it}=\log\varepsilon_{it}^2+1{\cdot}27$, $i=1,\ldots,N$, respectively, and ι is an $N\times 1$ vector of ones; compare the seemingly unrelated time series equation (SUTSE) models described in Harvey (1989, Chapter 8). Treating (9) as a Gaussian state-space model, QML estimators may be obtained by means of the Kalman filter. As in the univariate model, ξ_t and η_t are uncorrelated even if the original disturbances are correlated.

It is shown in Appendix B that the ij-th element of the covariance matrix of ξ_t, denoted Σ_ξ, is given by $(\pi^2/2)\rho_{ij}^*$, where $\rho_{ii}^*=1$ and

$$\rho_{ij}^*=\frac{2}{\pi^2}\sum_{n=1}^{\infty}\frac{(n-1)!}{(1/2)_n n}\rho_{ij}^{2n}, \qquad i\neq j,\ i,j=1,\ldots,N, \tag{10}$$

where $(x)_n=x(x+1)\ldots(x+n-1)$. Thus the absolute values of the unknown parameters in Σ_ε, namely the ρ_{ij}'s, the cross-correlations between different ε_{it}'s, may be estimated, but their signs may not, because the relevant information is lost when the observations are squared. However, estimates of the signs may be obtained by returning to the untransformed observations and noting that the sign of each of the pairs $\varepsilon_i\varepsilon_j$, $i,j=1,\ldots,N$, will be the same as the corresponding pair of observed values y_iy_j. Thus the sign of ρ_{ij} is estimated as positive if more than one-half of the pairs y_iy_j are positive.

1. If the state space form for the $\log y_{it}^2$'s were Gaussian, the conditional covariance between any two observations, y_{it} and y_{jt}, at time $t-1$, divided by their conditional standard deviations, would give the conditional correlation $\rho_{ij}\exp\{(2p_{ij,t|t-1}-p_{ii,t|t-1}-p_{jj,t|t-1})/8\}$, provided the ε_{it}'s and η_{it}'s are mutually independent. The terms $p_{ij,t|t-1}$ denote the ij-th elements of the covariance matrix of the filtered estimators of the h_{it}'s at time t, and since these are constant in the steady state, the conditional correlations are also constant.

4. COMMON FACTORS

In the K-factor GARCH model proposed by Engle (1987) the conditional covariance matrix depends on the conditional variances of K orthogonal linear combinations of y_t. Although the model can, in principle, be estimated by maximum likelihood, researchers often encounter computational difficulty with a large number of parameters. Engle, Ng and Rothschild (1990) suggest a simpler two-stage procedure. Bollerslev and Engle (1993) give conditions for covariance stationarity of K-factor GARCH models and show how multivariate IGARCH models allow for the possibility of co-persistence in variance. However, as in the univariate case, there is some ambiguity about what constitutes persistence.

An alternative multivariate model, which is not nested within multivariate GARCH, is the latent factor model of Diebold and Nerlove (1989). The model is a relatively parsimonious one in which the common movements in volatility are ascribed to a single unobserved latent factor subject to ARCH effects. However, this latent factor gives rise to similar common movements in the levels and for many purposes the levels and volatility effects need to be modelled separately.

Common factors can be incorporated in multivariate stochastic variance models very easily by following the literature on common factors in unobserved components time-series models; see Harvey (1989, Chapter 8, Section 5) for a review and Harvey and Stock (1988) for an application to U.S. data on income and consumption. We will concentrate on the case where there are persistent movements in volatility, modelled by a multivariate random walk. Thus (9) becomes

$$w_t = -1{\cdot}27\iota + \theta h_t + \bar{h} + \xi_t, \tag{11a}$$

$$h_t = h_{t-1} + \eta_t, \qquad \mathrm{Var}\,(\eta_t) = \Sigma_\eta, \tag{11b}$$

where θ is an $N\times k$ matrix of coefficients with $k \leqq N$, h_t and η_t are $k\times 1$ vectors, Σ_η is a $k\times k$ positive definite matrix and $\bar{h}$ is a $N\times 1$ vector in which the first k elements are zeroes while the last $N-k$ elements are unconstrained. The logarithm of variance for the i-th series is the i-th element of $\theta h_t + \bar{h}$. If $k<N$, the w_t's are co-integrated in the sense of Engle and Granger (1987). In the context of (11) this implies that there are $N-k$ linear combinations of the w_t's which are white noise.

As it stands model (11) is not identifiable. An identifiable model may be set up by requiring that the elements of θ are such that $\theta_{ij}=0$ for $j>i$, $i=1,\ldots,N$, $j=1,\ldots,k$, while Σ_η is an identity matrix. These restrictions are easily imposed, and the model may be estimated by QML using the Kalman filter to compute the prediction errors. Once this has been done, it may be worthwhile considering a rotation of the common factors to get a model with a more useful interpretation. If R is a $k\times k$ orthogonal matrix, the factors $h_t^* = Rh_t$ are still driven by mutually uncorrelated disturbances with unit variances, while the factor loading matrix becomes $\theta^* = \theta R'$.

The finite-sample properties of the QML estimator of model (11) have been studied by carrying out several Monte Carlo experiments. These are reported in Ruiz (1992) and confirm that the method works well for moderate sample sizes. The number of unknown parameters in θ is $(N-k)k + \frac{1}{2}k(k+1)$, while there are a further $\frac{1}{2}N(N-1)$ in Σ_ε. Numerical optimization must be carried out with respect to these unknown parameters. We used the quasi-Newton algorithm, EO4 AZF, in the NAG library.

5. EMPIRICAL APPLICATION: DAILY EXCHANGE RATES

In this section, the stochastic variance model is fitted to four exchange rates: Pound/Dollar, Deutschmark/Dollar, Yen/Dollar and Swiss-Franc/Dollar. The data consist of

TABLE 1.

Box–Ljung Q-statistics, based on ten lags, for daily exchange rates, p_t, of various currencies against the dollar

	$\Delta \log p_t$	$(\Delta \log p_t)^2$	$\log (\Delta \log p_t)^2$
Pound	11·19	128·25	45·47
DM	10·03	67·79	64·20
Yen	16·92	109·79	64·67
Swiss Franc	32·67	343·09	57·94

TABLE 2

Estimation results for univariate stochastic volatility models: (a) AR(1); (b) random walk

		\$/Pound	\$/DM	\$/Yen	\$/SF
(a)	$\hat{\varphi}$	0·9912	0·9646	0·9948	0·9575
		(0·0069)	(0·0206)	(0·0046)	(0·0024)
	$\hat{\sigma}_\eta^2$	0·0069	0·0312	0·0048	0·0459
		(0·0050)	(0·0219)	(0·0034)	(0·0291)
	γ	−0·0879	−0·3556	−0·0551	−0·4239
	$\log L$	−1212·82	−1232·26	−1272·64	−1288·51
(b)	$\hat{\sigma}_\eta^2$	0·0042	0·0161	0·0034	0·0194
		(0·0023)	(0·0063)	(0·0019)	(0·0072)
	$\log L$	−1214·02	−1237·58	−1273·46	−1294·22
	$Q(10)$	3·52	11·41	8·45	8·68

daily observations of weekdays close exchange rates from 1/10/81 to 28/6/85. The sample size is $T=946$. Table 1 shows Box–Ljung statistics for several transformations of the exchange rates. The chi-square 5% critical value for ten degrees of freedom is 18·3. With the possible exception of the Swiss Franc, the logarithms of the exchange rates appear to be random walks. The important point is that there is strong evidence of nonlinearity in the statistics for the squared differences and their logarithms.

Univariate models were fitted to the differences of the logarithms of each of the exchange rates, with the mean subtracted, that is

$$y_{it} = \Delta \log p_{it} - (\Sigma \Delta \log p_{it})/T, \qquad i=1, \ldots, N, \qquad t=1, \ldots, T.$$

Subtracting the mean ensures that there are no y_{it}'s identically equal to zero; this could create difficulties when logarithms of y_t^2 are taken. The QML estimates of the parameters φ and σ_η^2 in the stationary AR(1) volatility model, (3), are shown in Table 2. The estimates of φ are all close to one and the random walk specification, (5), fits almost as well. Asymptotic standard errors, based on Dunsmuir (1979, p. 502), are shown in parentheses, though it should be noted that they cannot be used to test whether σ_η^2 is significantly different from zero; see Harvey (1989, pp. 212–3). The Box–Ljung Q-statistics give no indication of residual serial correlation. Figure 1 shows the absolute values, $|y_t|$, for the Pound/Dollar series, together with the estimated standard deviation, $\exp(\frac{1}{2}\tilde{h}_{t|T})$, where $\tilde{h}_{t|T}$ is the MMSLE of the volatility level, h_t, as given by a smoothing algorithm.

The augmented Dickey–Fuller test applied to $\log y_t^2$, with nine lags and a constant included, rejects the hypothesis of a unit root at the 1% level for all the series; see Table 3. The significance point, for 500 observations, is −3·43 and so the rejection is quite decisive; using a smaller number of lags gave test statistics even further from the critical value. However, the reliability of unit root tests in this situation is questionable. The

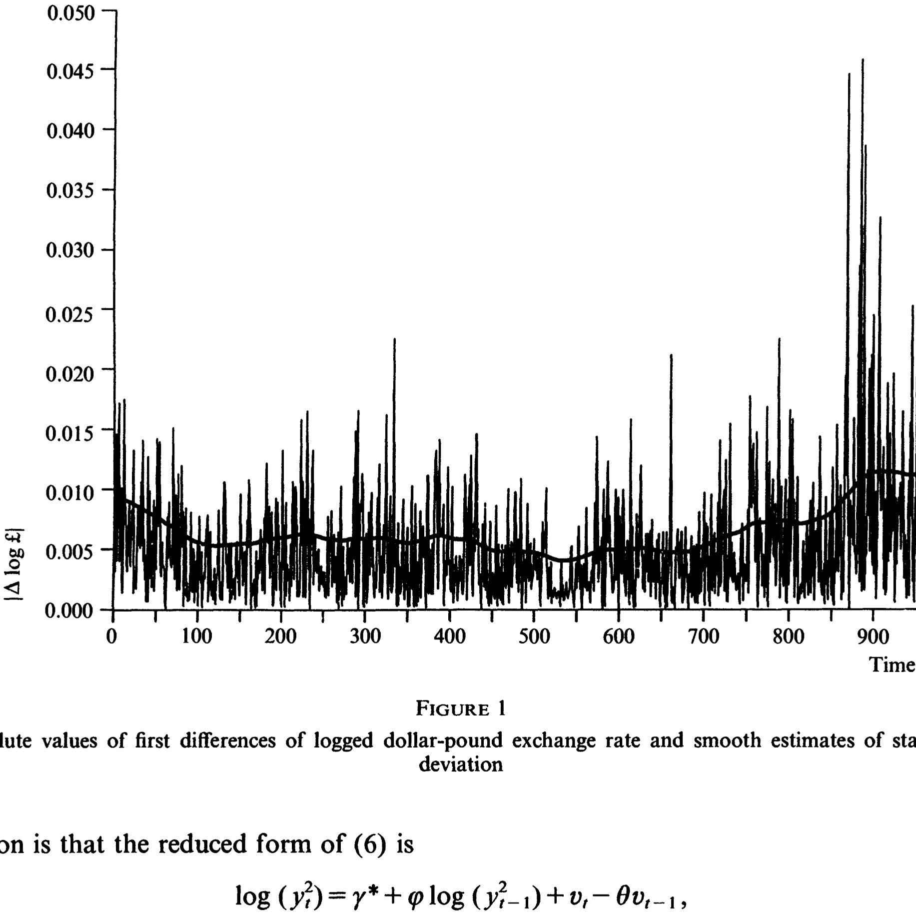

FIGURE 1

Absolute values of first differences of logged dollar-pound exchange rate and smooth estimates of standard deviation

reason is that the reduced form of (6) is

$$\log(y_t^2) = \gamma^* + \varphi \log(y_{t-1}^2) + v_t - \theta v_{t-1}, \tag{12}$$

where v_t is white noise and $\gamma^* = (\gamma - 1{\cdot}27)/(1-\varphi)$. Since the variance of ξ_t typically dominates the variance of η_t, the parameter θ will be close to unity for values of φ close to one. For example for the Dollar/Pound exchange rate, where the estimated φ value is 0·99, the implied θ is −0·97. As shown in Pantula (1991) and Schwert (1989), when the moving-average parameter is very close to one, unit root tests reject the null hypothesis of a unit root too often since the model is difficult to distinguish from white noise.

TABLE 3

Augmented Dickey-Fuller test statistics for the logarithms of squared differences of logarithms of daily exchange rates

	\$/Pound	\$/DM	\$/Yen	\$/SF
'*t*-stat'	−7·42	−7·50	−7·63	−7·44

Since unit root tests based on autoregressive approximations are unreliable in the present context, there is little point in trying to determine the number of common trends on the basis of co-integration tests such as the one described in Johansen (1988). Instead we estimate the unrestricted multivariated local level model, (9), and make a judgement as to the number of possible common trends on the basis of a principal component analysis of the estimate of Σ_η.

QML estimation of (9), with the diagonal elements of the matrix Σ_ξ set to $\pi^2/2$, gives:

$$\hat{\Sigma}_\xi = \frac{\pi^2}{2}\begin{bmatrix} 1\cdot00 & 0\cdot404 & 0\cdot278 & 0\cdot347 \\ & 1\cdot000 & 0\cdot400 & 0\cdot541 \\ & & 1\cdot000 & 0\cdot362 \\ & & & 1\cdot000 \end{bmatrix}$$

and

$$\hat{\Sigma}_\eta = 10^{-3}\begin{bmatrix} 9\cdot65 & 11\cdot42 & 3\cdot97 & 12\cdot07 \\ & 20\cdot43 & 5\cdot44 & 21\cdot09 \\ & & 5\cdot45 & 7\cdot08 \\ & & & 22\cdot31 \end{bmatrix}$$

the value of the maximized quasi-log-likelihood for the multivariate model, $-4621\cdot64$, is substantially greater than the sum of $-4963\cdot06$ for the four univariate models. The number of additional parameters in the multivariate model is 12, and so the value of 682·8 taken by the (quasi-) likelihood ratio test statistic, is highly significant if judged against a χ^2_{12} distribution.

From (10), the implied covariance (correlation) matrix for ε_t is

$$\hat{\Sigma}_\varepsilon = \begin{bmatrix} 1\cdot00 & 0\cdot84 & 0\cdot74 & 0\cdot80 \\ & 1\cdot00 & 0\cdot84 & 0\cdot92 \\ & & 1\cdot00 & 0\cdot81 \\ & & & 1\cdot00 \end{bmatrix}$$

Estimating the signs of the cross correlations in Σ_ε from the signs of pairs of y_t's indicated they were all positive. The correlation matrix corresponding to the estimate of Σ_η is

$$\text{Corr}\,(\hat{\eta}) = \begin{bmatrix} 1\cdot00 & 0\cdot81 & 0\cdot55 & 0\cdot82 \\ & 1\cdot00 & 0\cdot51 & 0\cdot99 \\ & & 1\cdot00 & 0\cdot64 \\ & & & 1\cdot00 \end{bmatrix}$$

It is interesting to note that the correlations between the elements of ε_t are uniformly high for all four exchange rates, while for η_t the correlations involving the Yen are much lower than the European currencies.

The results of a principal components analysis of Σ_η and its correlation matrix appear in Table 4. The units of measurement are not relevant since logarithms have been taken, but differences appear in the results for the covariance and correlation matrices, primarily because the Yen shows much less variation than the other three exchange rates. The two first components account for 94% or 95% of the total variance of the disturbance η_{it}, $i = 1, 2, 3, 4$. The second component is relatively more important when the correlation matrix is analysed, with a fairly high loading on the Yen. Table 5 shows the first two eigenvectors multiplied by the square roots of the corresponding eigenvalues. For the analysis of the

TABLE 4

Principal components analysis of (a) $\hat{\Sigma}_\eta$ *and* (b) *corresponding correlation matrix*

(a)	Eigenvalues	0·0515	0·0037	0·0027	$1{\cdot}4281\times 10^{-6}$
	Eigenvectors	0·3781	−0·1443	0·9144	−0·0139
		0·6205	0·3749	−0·1873	0·6629
		0·2066	−0·9156	−0·2260	0·2605
		0·6553	0·0170	−0·2789	−0·7018
	Percentage of variance	88·98	6·43	4·58	0·00
2b)	Eigenvalues	3·1933	0·5703	0·2363	$7{\cdot}02\times 10^{-5}$
	Eigenvectors	0·5033	0·2102	0·8381	−0·0095
		0·5290	0·3476	−0·3973	0·6644
		0·4097	−0·9020	−0·0182	0·1348
		0·5468	0·1460	−0·3733	−0·7350
	Percentage of variance	79·83	14·26	5·91	0·00

TABLE 5

Principal components analysis: First two eigenvectors multiplied by square roots of corresponding eigenvalues for (a) *correlation matrix, and* (b) *covariance matrix*

Series	(a)		(b)	
1	0·8994	0·1587	0·0858	−0·0088
2	0·9453	0·2625	0·1408	0·0229
3	0·7321	−0·6812	0·0469	−0·0558
4	0·9771	0·1103	0·1487	0·0010

correlation matrix these figures give the correlations between the principal components and the disturbances η_{it}, $i=1, 2, 3, 4$. In this case, the first component can perhaps be interpreted as a general underlying factor, strongly correlated with the European exchange rates, but less so with the Yen, while the second component is correlated most strongly with the Yen. The loadings for the first component in the analysis of the covariance matrix invite a similar interpretation, but the Yen is not dominant in the second component.

In principal components analysis, the covariance matrix is decomposed as $\hat{\Sigma}_\eta = EDE'$, where E is the matrix of eigenvectors and D is a diagonal matrix of eigenvalues. The principal components, $E'w_t$, have covariance matrix D. Noting that $ED^{1/2}D^{-1/2}E'$ is an identity matrix, the model in (9) may be written as

$$w_t = -1{\cdot}27i + \theta h_t^* + \xi_t \tag{13a}$$

$$h_t^* = h_{t-1}^* + \eta_t^*, \qquad \mathrm{Var}(\eta_t^*) = I \tag{13b}$$

where $h_t^* = D^{-1/2}E'h_t$ and $\theta = ED^{1/2}$. This provides a useful link with model (11) when $k = N$, the necessary restrictions on θ coming from the properties of standardised eigenvectors rather than by setting elements above the leading diagonal to zero. If the estimate of Σ_η were of rank k, then θ would be an $N\times k$ matrix. Note that in the present application, the first two columns of θ are given by the entries in Table 5.

The principal components analysis suggests that two factors might be enough to account for the movements in volatility. Estimating (11) with $k=2$, and the restrictions $\theta_{12}=0$ and $\Sigma_\eta = I$ gives:

$$\log(\hat{y}_{1t}^2) = -1{\cdot}27 + 0{\cdot}108\,\hat{h}_{1t}$$

$$\log(\hat{y}_{2t}^2) = -1{\cdot}27 + 0{\cdot}102\,\hat{h}_{1t} + 0{\cdot}014\,\hat{h}_{2t}$$

$$\log(\hat{y}_{3t}^2) = -1\cdot27 + 0\cdot016\,\hat{h}_{1t} + 0\cdot054\,\hat{h}_{2t} - 7\cdot42$$
$$\log(\hat{y}_{4t}^2) = -1\cdot27 + 0\cdot095\,\hat{h}_{1t} + 0\cdot023\,\hat{h}_{2t} - 1\cdot38$$

and

$$\hat{\Sigma}_{\xi} = \frac{\pi^2}{2}\begin{bmatrix} 1\cdot00 & 0\cdot382 & 0\cdot271 & 0\cdot334 \\ & 1\cdot00 & 0\cdot390 & 0\cdot539 \\ & & 1\cdot00 & 0\cdot358 \\ & & & 1\cdot00 \end{bmatrix}$$

with a quasi-log-likelihood of $-4626\cdot48$. The implied correlation matrix for ε_t is

$$\hat{\Sigma}_{\varepsilon} = \begin{bmatrix} 1\cdot00 & 0\cdot83 & 0\cdot73 & 0\cdot79 \\ & 1\cdot00 & 0\cdot83 & 0\cdot92 \\ & & 1\cdot00 & 0\cdot81 \\ & & & 1\cdot00 \end{bmatrix}.$$

Again, estimating the signs of the cross-correlations in Σ_ε from the signs of pairs of y_t's indicates that they are all positive.

Factor rotation was carried out using the orthogonal matrix

$$R = \begin{bmatrix} \cos\lambda & -\sin\lambda \\ \sin\lambda & \cos\lambda \end{bmatrix}$$

and a graphical method; see Schuessler (1971). For clockwise rotation, setting the angle, λ, to $16\cdot23°$ gives a loading of zero for the first factor on the third series, the Yen; see Figure 2 and Table 6. Setting the angle to $352\cdot48°$ gives a loading of zero for this second factor on the DM and very small loadings on the other two European currencies. The first rotation therefore has the dominant factor, the first common trend h_{1t}, related only to the European exchange rates, while the second common trend, h_{2t}, is a general trend which underlies the volatility of all the exchange rates including the Yen. In the second rotation, which is actually quite close to the original when the movement is in an anti-clockwise direction, the first common trend affects the European exchange rates to a similar extent, but leads to smaller movements in the Yen. The second common trend has its effect almost exclusively on the Yen; compare the results for the principal components analysis of the correlation matrix as given in Table 4. The message in the two rotations is essentially the same and which one is adopted is really a matter of taste. The standard deviations implied by the two common trends for the second rotation, $\exp(\frac{1}{2}\tilde{h}_{jt|T})$, $j=$

TABLE 6

Loadings after rotations

	$\lambda=16\cdot23^0$		$\lambda=352\cdot48^0$	
Pound/Dollar	0·103	0·030	0·107	−0·014
DM/Dollar	0·095	0·042	0·103	0
Yen/Dollar	0	0·056	0·023	0·051
SF/Dollar	0·085	0·048	0·097	0·010

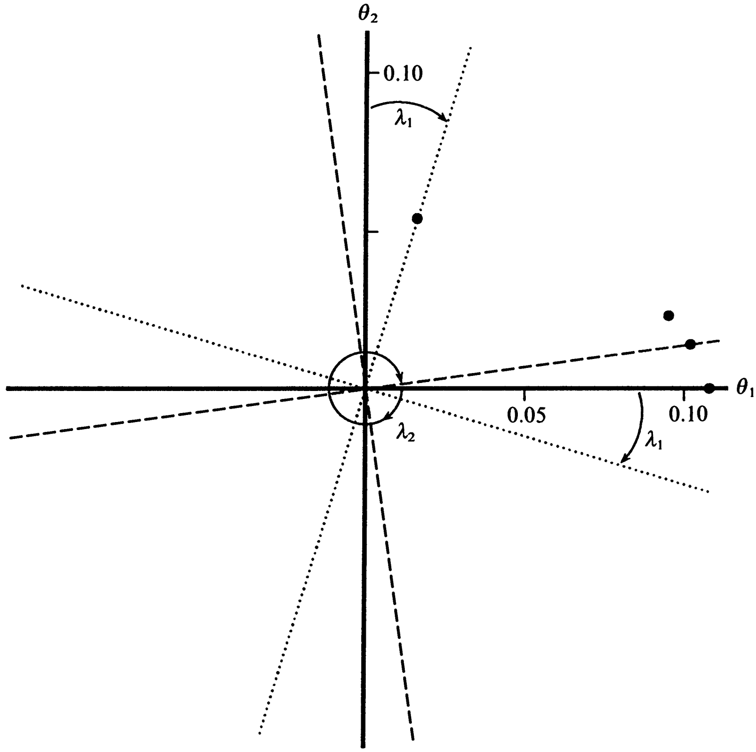

FIGURE 2
Factor rotation

1, 2, are plotted in Figure 3. The standard deviations estimated from the univariate models for the Deutschmark and Yen are shown for comparison.

6. HEAVY-TAILED DISTRIBUTIONS

The GARCH model may be generalized by letting ε_t have a Student t-distribution; see Bollerslev (1987). This is important because the kurtosis in many financial series is greater than the kurtosis which results from incorporating conditional heteroscedasticity into a Gaussian process. A similar generalization is possible for the SV model. Once again it can be shown that when h_t is stationary, y_t is white noise and it follows immediately from the properties of the t-distribution that the formula for the unconditional variance generalizes to $\{\nu/(\nu-2)\}\exp(\gamma_h+\frac{1}{2}\sigma_h^2)$.

Let ε_t in (1) be a t-variable written as

$$\varepsilon_t=\zeta_t/\kappa_t^{1/2}, \qquad t=1,\ldots,T, \tag{14}$$

where ζ_t is a standard normal variate and $\nu\kappa_t$ is distributed, independently of ζ_t, as a χ^2 with ν degrees of freedom. Thus

$$\log\varepsilon_t^2=\log\zeta_t^2-\log\kappa_t, \tag{15}$$

and it follows from Abramovitz and Stegun (1970, p. 943) that the mean and variance of

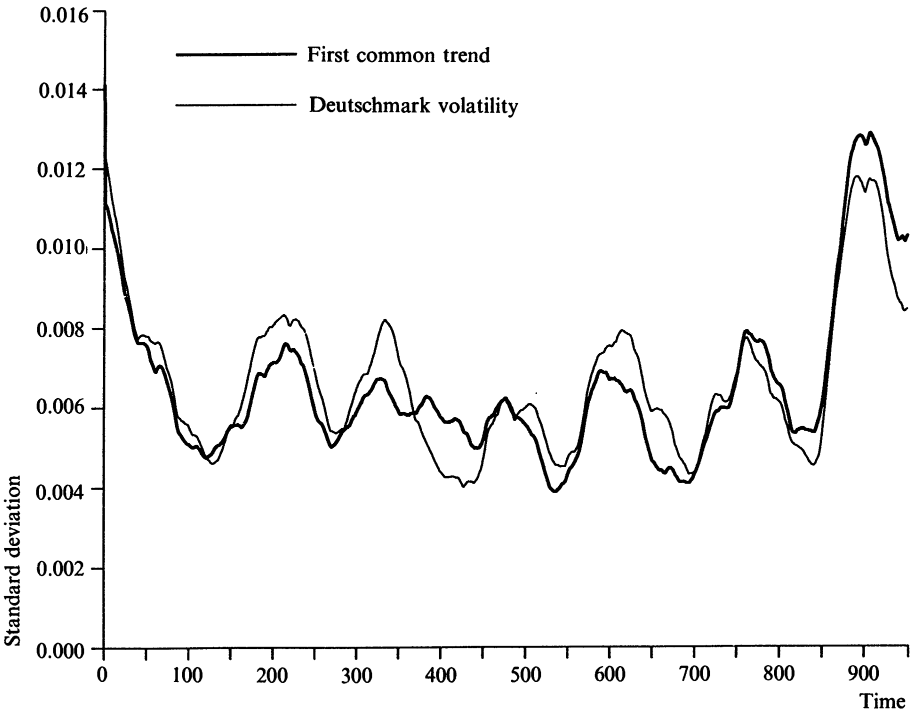

FIGURE 3(a)

Smooth estimates of standard deviation for the first common trend with the corresponding estimates for the Deutschmark

$\log \kappa_t$ are $\psi'(\nu/2) - \log(\nu/2)$ and $\psi'(\nu/2)$ respectively, where $\psi(\cdot)$ and $\psi'(\cdot)$ are the digamma and trigamma functions respectively. The ACF of $\log y_t^2$ has the same form as before.

The state-space model corresponding to (6) can be estimated for a given value of ν. Alternatively the variance of $\log \varepsilon_t^2$ can be treated as an additional unknown parameter. In both cases the asymptotic theory in Dunsmuir (1979, p. 502) applies since the required moments of $\log \varepsilon_t^2$ exist even when ν is one. Leaving the distribution of ε_t unspecified means that γ_h is not identified since the expected value of $\log \varepsilon_t^2$ is unknown. Similarly when h_t follows a random walk its estimated values will include the expectation of $\log \varepsilon_t^2$. Thus the level of volatility is not determined. However, if ε_t is assumed to have a t-distribution, the estimated variance of $\log \varepsilon_t^2$ implies a value of ν when set to $4{\cdot}93 + \psi'(\nu/2)$, and this in turn enables the expectation of $\log \varepsilon_t^2$ to be calculated. In the exchange rate application, the unrestricted estimates of the variance of $\log \varepsilon_t^2$ imply that the distribution of ε_t is normal for the Pound and Deutschmark, that is ν is infinity, while for the Yen and Swiss Franc ν is approximately six.

The generalization to the multivariate model can be made by assuming that (14) holds for $i = 1, \ldots, N$ with the ζ_{it}'s following a multivariate normal distribution, with a correlation matrix as specified for ε_t in (7), but with the κ_{it}'s mutually independent. The covariance matrix of the vector of $\log \varepsilon_{it}^2$ variables, Σ_ξ, is the sum of the covariance matrix of the $\log \zeta_{it}^2$ variables, defined as in (10), and a diagonal matrix in which the i-th diagonal element is the variance of $\log \kappa_{it}$. Each diagonal element in the covariance matrix of ε_t, Σ_ε, is equal to the variance of the corresponding t-distribution, that is

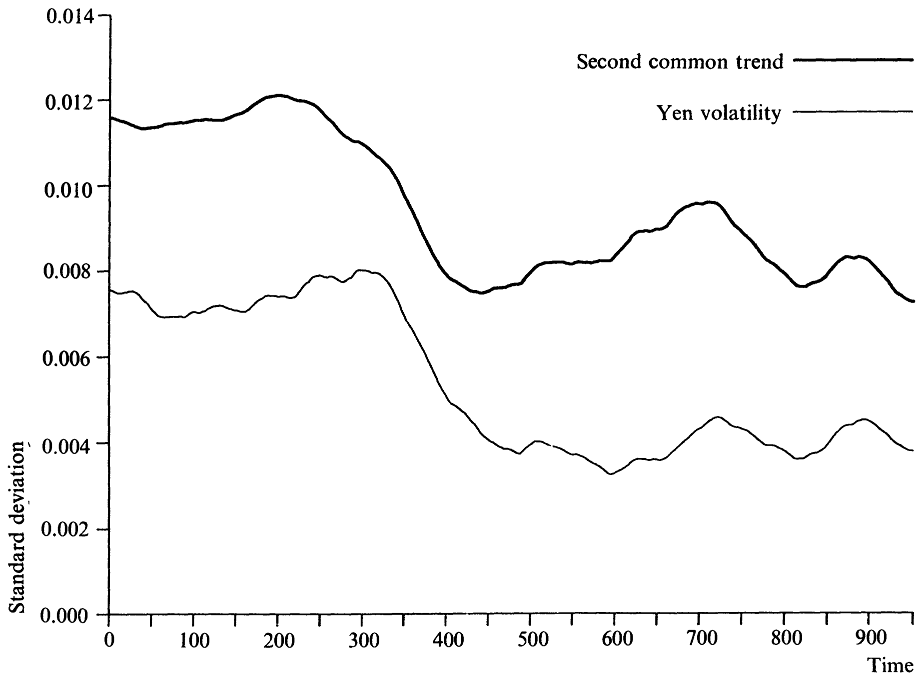

FIGURE 3(b)

Smooth estimates of standard deviation for the second common trend with the corresponding estimates for the Yen

Var $(\varepsilon_{it}) = \nu_i/(\nu_i - 2)$ for $\nu_i > 2$ and $i = 1, \ldots, N$. As regards the off diagonal elements

$$E(\varepsilon_{it}\varepsilon_{jt}) = E(\kappa_i^{-1/2})E(\kappa_j^{-1/2})E(\zeta_i\zeta_j), \qquad i \neq j, \qquad i, j = 1, \ldots, N.$$

The last term is obtained from the corresponding covariance of $\log \zeta_i^2$ and $\log \zeta_j^2$ using (10), while

$$E(\kappa_i^{-1/2}) = \nu_i^{1/2} 2^{-1/2} \Gamma\{(\nu_i - 1)/2\}/\Gamma(\nu_i/2), \qquad i = 1, \ldots, N.$$

Fitting the above multivariate model gave the following results

$$\hat{\Sigma}_\xi = \frac{\pi^2}{2}\begin{bmatrix} 1\cdot 00 & 0\cdot 411 & 0\cdot 303 & 0\cdot 380 \\ & 1\cdot 00 & 0\cdot 434 & 0\cdot 586 \\ & & 1\cdot 00 & 0\cdot 419 \\ & & & 1\cdot 000 \end{bmatrix} + \begin{bmatrix} 0\cdot 000 & & & \\ & 0\cdot 000 & & \\ & & 0\cdot 406 & \\ & & & 0\cdot 510 \end{bmatrix}$$

and

$$\hat{\Sigma}_\eta = 10^{-3}\begin{bmatrix} 8\cdot 69 & 10\cdot 25 & 2\cdot 92 & 10\cdot 24 \\ & 19\cdot 16 & 4\cdot 07 & 18\cdot 82 \\ & & 4\cdot 02 & 5\cdot 21 \\ & & & 18\cdot 96 \end{bmatrix}$$

The maximized quasi-log-likelihood, −4618·06, is slightly higher than for the Gaussian model reported in Section 5. The implied degrees of freedom for the Yen and Swiss Franc are 5·86 and 4·84 respectively.

The correlation matrix for ε_t is

$$\Sigma_\varepsilon = \begin{bmatrix} 1\cdot00 & 0\cdot85 & 0\cdot88 & 0\cdot99 \\ & 1\cdot00 & 0\cdot99 & 1\cdot12 \\ & & 1\cdot52 & 1\cdot18 \\ & & & 1\cdot70 \end{bmatrix}.$$

The covariance matrix of η_t is not very different from the one reported for the Gaussian model, and the same is true of the correlation matrix

$$\mathrm{Corr}\,(\hat{\eta}) = \begin{bmatrix} 1\cdot00 & 0\cdot79 & 0\cdot49 & 0\cdot80 \\ & 1\cdot00 & 0\cdot46 & 0\cdot99 \\ & & 1\cdot00 & 0\cdot60 \\ & & & 1\cdot00 \end{bmatrix}.$$

As a result the common trends, and the implied groupings of exchange rates, are similar.

7. CONCLUSION

The multivariate stochastic variance model has a natural interpretation and is relatively parsimonious. The parameters can be estimated without too much difficulty by a quasi-maximum likelihood approach, and the movements in variance can be estimated by smoothing. The extension to heavier tailed distributions can be carried out very easily using the *t*-distribution.[2] The model fits well to exchange rates, and is able to capture common movements in volatility. The volatility in the three European exchange rates depends primarily on one factor. This factor affects the Yen to a much lesser extent, and the Yen is primarily affected by a second factor. Other rotations offer a slightly different interpretation but the special behaviour of the Yen is always apparent.

APPENDIX A. UNCORRELATEDNESS OF VARIABLES AFTER TRANSFORMATION

Consider two random variables, ε and η, which may be dependent. Assume $E(\eta)=0$ and let $h(\cdot)$ be an even function such that $E[h(\varepsilon)]$ exists. If the covariance between η and $h(\varepsilon)$ exists, it is zero under the following conditions:

A.1 The density of ε, $f(\varepsilon)$, is symmetric.
A.2 $E(\eta \mid \varepsilon)$ is an odd function of ε.

The result follows because

$$\mathrm{Cov}\,(\eta, h(\varepsilon)) = E[\eta(h(\varepsilon) - E[h(\varepsilon)])] = E[\eta h(\varepsilon)] = E[E(\eta \mid \varepsilon)h(\varepsilon)].$$

and, under A.2, $E[E(\eta \mid \varepsilon)h(\varepsilon)]$ is an odd function of ε, and so given A.1 its expected value is zero.

2. The STAMP package can be used to carry out estimation by QML for univariate models. A multivariate version is currently being developed. Further information can be obtained by writing to the first author at LSE.

In the application here, $h(\varepsilon)=\log\varepsilon^2$ is an even function, and if ε and η are zero mean bivariate normal, conditions A.1 and A.2 are satisfied, and so

$$\operatorname{Cov}(\eta, \log\varepsilon^2)=0.$$

When ε has a Student t-distribution with v degrees of freedom, it can be written as $\varepsilon=\zeta\kappa^{-1/2}$, where $v\kappa$ is distributed independently of ζ as χ^2_v. If ζ and η are bivariate normal, then

$$E(\eta\log\varepsilon^2)=E[\eta(\log\zeta^2-\log\kappa)]=E[\eta\log\zeta^2]-E(\eta)E(\log\kappa)=0$$

and so η and $\log\varepsilon^2$ are again uncorrelated. Note that the result holds even if ε is Cauchy distributed ($v=1$), since although the mean of ε does not exist in this case, $E[\log\varepsilon^2]$ does exist and in fact is zero.

Conditions A.1 and A.2 are satisfied if the joint distribution of ε and η satisfies the symmetry condition:

A.3 $g(\varepsilon, \eta)=g(-\varepsilon, -\eta)$.

This follows because A.3 implies A.1 since

$$f(\varepsilon)=\int_{-\infty}^{\infty} g(\varepsilon, \eta)d\eta=\int_{-\infty}^{\infty} g(-\varepsilon, -\eta)d\eta=\int_{-\infty}^{\infty} g(-\varepsilon, \eta)d\eta=f(-\varepsilon).$$

while A.3 implies A.2 because

$$E(\eta\mid\varepsilon)=\int_{-\infty}^{\infty}\eta\frac{g(\varepsilon,\eta)}{f(\varepsilon)}d\eta=\frac{1}{f(\varepsilon)}\int_{-\infty}^{\infty}\eta g(-\varepsilon, -\eta)d\eta$$
$$=\frac{-1}{f(-\varepsilon)}\int_{-\infty}^{\infty}\eta g(-\varepsilon, \eta)d\eta=-E(\eta\mid-\varepsilon).$$

APPENDIX B. CORRELATIONS BETWEEN TRANSFORMATIONS OF STANDARD NORMAL VARIABLES

In this appendex we derive the expression for Corr ($\log\varepsilon_1^2$, $\log\varepsilon_2^2$) where ε_1 and ε_2 are bivariate standard normal variables with correlation coefficient ρ.

Define $u=(\varepsilon_1^2)^{1/2}$ and $v=(\varepsilon_2^2)^{1/2}$. Johnson and Kotz (1972) give the following expression for the moments of u and v

$$E(u^r v^s)=\frac{1}{(\Gamma(1/2))^2}\left[2^{(r+s)/2}\Gamma\left(\frac{1+r}{2}\right)\Gamma\left(\frac{1+s}{2}\right)F\left(\frac{r}{2},\frac{2}{2};\frac{1}{2};\rho^2\right)\right] \tag{B.1}$$

where $F(a, b; c; z)$ is the hypergeometric function given by

$$F(a,b;c;z)=\sum_{n=0}^{\infty}\frac{(a)_n(b)_n}{(c)_n}\frac{z^n}{n!}=1+\frac{ab}{\gamma}z+\frac{a(a+1)b(b+1)}{2!\gamma(\gamma+1)}z^2+\cdots \tag{B.2}$$

where $(a)_n=\Gamma(a+n)/\Gamma(a)$.

The moment generating function of $\log\varepsilon_1^2$ and $\log\varepsilon_2^2$ is given by

$$m(t_1, t_2)=E[\exp\{t_1\log\varepsilon_1^2+t_2\log\varepsilon_2^2\}]=E[\varepsilon_1^{2t_1}\varepsilon_2^{2t_2}]. \tag{B.3}$$

Using (B.1) in (B.3) and taking logarithms yields

$$\log m(t_1, t_2)=(t_1+t_2)\log(2)+\log\Gamma((1/2)+t_1)+\log\Gamma((1/2)+t_2)$$
$$+\log F(-t_1, -t_2; 1/2; \rho^2)-2\log\Gamma(1/2) \tag{B.4}$$

and, therefore,

$$\frac{\partial^2\log m(t_1,t_2)}{\partial t_1\partial t_2}=\frac{1}{F}\frac{\partial^2 F}{\partial t_1\partial t_2}-\frac{1}{F^2}\frac{\partial F}{\partial t_1}\frac{\partial F}{\partial t_2}. \tag{B.5}$$

To find the covariance between $\log \varepsilon_1^2$ and $\log \varepsilon_2^2$, expression (B.5) has to be evaluated at $t_1 = t_2 = 0$. Given the expression for the hypergeometric function in (B.2), it is easy to see that at $t_1 = t_2 = 0$

$$F(0, 0; 1/2; \rho^2) = 1 \tag{B.6}$$

$$\frac{\partial F}{\partial t_1} = \frac{\partial F}{\partial t_2} = 0 \tag{B.7}$$

and

$$\frac{\partial^2 F}{\partial t_1 \partial t_2} = \sum_{n=1}^{\infty} \frac{((n-1)!)^2}{(1/2)_n} \frac{\rho^{2n}}{n!} = \sum_{n=1}^{\infty} \frac{(n-1)!}{(1/2)_n n} \rho^{2n}. \tag{B.8}$$

Substituting (B.6), (B.7) and (B.8) into (B.5), we get

$$\text{Cov}\,(\log \varepsilon_1^2, \log \varepsilon_2^2) = \sum_{n=1}^{\infty} \frac{(n-1)!}{(1/2)_n n} \rho^{2n} \tag{B.9}$$

The variance of $\log \varepsilon_i^2$ is given by $\pi^2/2$ for $i = 1, 2$ and therefore the correlation is as in expression (10) in the text.

Acknowledgements. A preliminary version of this paper was presented at the EC^2 conference in Rotterdam in December, 1991, and at seminars at LSE, Oxford, INSEE, Florence, Bilbao, Bank of Spain, Pompeu Fabra, Princeton, Harvard–MIT and Warwick. We are grateful to the participants, and to Javier Hidalgo, Alain Monfort, Enrique Sentana and Stephen Taylor for helpful comments. The referees also made a number of helpful comments and suggestions. All errors are our responsbility.

Support from the ESRC, as part of the project "Estimation and Diagnostic Testing of Unobserved Components Models", grant No. R000 23 3574, is gratefully acknowledged.

REFERENCES

ABRAMOWITZ, M. and STEGUN, N. C. (1970) *Handbook of Mathematical Functions* (New York: Dover Publications Inc.).

BAILLIE, R. T. and BOLLERSLEV, T. (1990), "A Multivariate Generalized ARCH Approach to Modeling Risk Premia in Forward Foreign Exchange Rate Markets", *Journal of International Money and Finance*, **9**, 309–324.

BOLLERSLEV, T. (1986), "Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics*, **51**, 307–327.

BOLLERSLEV, T. (1987), "A Conditional Heteroskedastic Time Series Model for Speculative Prices and Rates of Return", *Review of Economics and Statistics*, **69**, 542–547.

BOLLERSLEV, T. (1990), "Modelling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized ARCH Approach", *Review of Economics and Statistics*, **72**, 498–505.

BOLLERSLEV, T. and ENGLE, R. F. (1993), "Common Persistence in Conditional Variances", *Econometrica*, **61**, 167–186.

BOLLERSLEV, T., ENGLE, R. F. and WOOLDRIDGE, J. M. (1988), "A Capital Asset Pricing Model with Time Varying Covariances", *Journal of Political Economy*, **96**, 116–131.

BOLLERSLEV, T., CHOU, R. Y. and KRONER, K. F. (1992), "ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence", *Journal of Econometrics*, **52**, 5–59.

CHESNEY, M. and SCOTT, L. (1989), "Pricing European Currency Options: A Comparison of the Modified Black–Scholes Model and a Random Variance Model", *Journal of Financial and Quantitative Analysis*, **24**, 267–284.

DASSIOS, A. (1992), "Asymptotic Approximations to Stochastic Variance Models" (mimeo, LSE).

DIEBOLD, F. X. and NERLOVE, M. (1989), "The Dynamics of Exchange Rate Volatility: A Multivariate Latent ARCH Model", *Journal of Applied Econometrics*, **4**, 1–21.

DUNSMUIR, W. (1979), "A Central Limit Theorem for Parameter Estimation in Stationary Vector Time Series and Its Applications to Models for a Signal Observed with Noise", *Annals of Statistics*, **7**, 490–506.

ENGLE, R. F. (1982), "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation", *Econometrica*, **50**, 987–1007.

ENGLE, R. F. (1987), "Multivariate GARCH with Factor Structures-Cointegration in Variance" (unpublished paper, UCSD).

ENGLE, R. F. and BOLLERSLEV, T. (1986), "Modelling the Persistence of Conditional Variances", *Econometric Reviews*, **5**, 1–50.

ENGLE, R. F. and GRANGER, C. W. J. (1987), "Co-integration and Error Correction: Representation, Estimation and Testing", *Econometrica*, **55**, 251–276.

ENGLE, R. F., NG, V. K. and ROTHSCHILD, M. (1990), "Asset Pricing with a Factor ARCH Covariance Structure: Empirical Estimates for Treasury Bills", *Journal of Econometrics*, **45**, 213–238.

HARVEY, A. C. (1989), *Forecasting, Structural Time Series Models and the Kalman Filter* (Cambridge: Cambridge University Press).

HARVEY, A. C. and STOCK, J. H. (1988), "Continuous Time Autoregressive Models with Common Stochastic Trends", *Journal of Economic Dynamics and Control*, **12**, 365–384.

HULL, J. and WHITE, A. (1987), "Hedging the Risks from Writing Foreign Currency Options", *Journal of International Money and Finance*, **6**, 131–152.

JOHANSEN, S. (1988), "Statistical Analysis of Cointegration Vectors", *Journal of Economic Dynamics and Control*, **12**, 231–254.

JOHNSON, N. L. and KOTZ, S. (1972) *Distributions in Statistics: Continuous Multivariate Distributions* (New York: John Wiley and Sons Inc.).

MELINO, A. and TURNBULL, S. M. (1990), "Pricing Foreign Currency Options with Stochastic Volatility", *Journal of Econometrics*, **45**, 239–265.

NELSON, D. B. (1988), "The Time Series Behaviour of Stock Market Volatility and Returns" (Unpublished Ph.D Dissertation, Massachusetts Institute of Technology).

NELSON, D. B (1990), "Stationarity and Persistence in the GARCH(1, 1) Model", *Econometric Theory*, **6**, 318–334.

NELSON, D. B. (1991), "Conditional Heteroscedasticity in Asset Returns: A New Approach", *Econometrica*, **59**, 347–370.

NELSON, D. B. and CAO, C. Q. (1992), "Inequality Constraints in the Univariate GARCH Model", *Journal of Business and Economic Statistics*, **10**, 229–235.

PANTULA, S. G. (1991), "Asymptotic Distributions of Unit-Root Tests when the Process is Nearly Stationary", *Journal of Business and Economic Statistics*, **9**, 63–71.

RUIZ, E. (1992), "Heteroscedasticity in Financial Time Series" (Unpublished Ph.D Dissertation, University of London).

RUIZ, E. (1994), "Quasi-Maximum Likelihood Estimation of Stochastic Volatility Models", *Journal of Econometrics* (forthcoming).

SCHUESSLER, K. (1971) *Analysing Social Data: A Statistical Orientation* (Boston: Houghton Mifflin Company).

SCHWERT, G. W. (1989), "Tests for Unit Roots: A Monte Carlo Investigation", *Journal of Business and Economic Statistics*, **7**, 147–160.

SCHWERT, G. W. and SEGUIN, P. J. (1990), "Heteroskedasticity in Stock Returns", *Journal of Finance*, **45**, 1129–1155.

SCOTT, L. O. (1987), "Option Pricing when the Variance Changes Randomly. Theory, Estimation and an Application", *Journal of Financial and Quantitative Analysis*, **22**, 419–438.

TAYLOR, S. J. (1986), *Modelling Financial Time Series* (Chichester: John Wiley).

TAYLOR, S. J. (1993), "Modelling Stochastic Volatility", *Mathematical Finance* (forthcoming).