Mathematical Finance, Vol. 4, No. 2 (April 1994), 183–204

MODELING STOCHASTIC VOLATILITY: A REVIEW AND COMPARATIVE STUDY

STEPHEN J. TAYLOR [1]

Department of Accounting and Finance, The Management School, Lancaster University, England

Diffusion models for volatility have been used to price options while ARCH models predominate in descriptive studies of asset volatility. This paper compares a discrete-time approximation of a popular diffusion model with ARCH models. These volatility models have many similarities but the models make different assumptions about how the magnitude of price responses to information alters volatility and the amount of subsequent information. Several volatility models are estimated for daily DM/$ exchange rates from 1978 to 1990.

KEY WORDS: conditional heteroskedasticity, exchange rates, stochastic volatility, volatility persistence

1. INTRODUCTION

Volatility changes occur for all classes of assets and have been reported in numerous stock, currency, and commodity studies. Important issues in volatility research include, first, the implications for option pricing, second, volatility estimation, and, third, whether volatility shocks persist indefinitely. These issues have been addressed using different models and methodologies. First, the finance literature contains diffusion models for a *stochastic volatility* variable which permit the rigorous valuation of options for such models (e.g., Hull and White 1987a; Chesney and Scott 1989; Heston 1993). The diffusion models have not been motivated by data analysis, and inferences about any mean reversion in volatility have been drawn primarily from studies of implied volatilities (e.g., Merville and Pieptea 1989, Stein 1989). Second, a very substantial econometric literature has focused on discrete-time ARCH models displaying *conditional heteroskedasticity,* following the path-breaking paper by Engle (1982). Maximum likelihood estimates for ARCH models permit a direct evaluation of the mean reversion issue.

The motivation for this paper is the opinion that an understanding of both ARCH and other models for volatility is more beneficial than knowledge of only one way to model volatility. For example, it is possible to numerically value options without recourse to ARCH methods (Chesney and Scott 1989; Melino and Turnbull 1990) or to do this solely using ARCH methods (Engle and Mustafa 1992; Duan 1993), but a methodology which takes advantage of both non-ARCH and ARCH methods provides more satisfactory results. Details are provided in Section 7.

[1] I thank all those who have helped me to clarify the arguments presented here, including participants at the 1990 European Finance Association meeting, the 1992 French Finance Association meeting, and seminars held at the European Institute for Advanced Studies in Management, Lancaster University, Monash University, Reading University, Stockholm School of Economics, and Warwick University.

Manuscript received April 1993; final revision received October 1993.

This paper reviews, compares, and estimates the volatility models studied in options and econometric research. Nelson (1990) and Bollerslev, Chou, and Kroner (1992) have respectively published important theoretical results and an excellent review of ARCH literature. This paper, however, provides a more comprehensive comparison of the various volatility models deserving serious consideration. Four models are considered in detail. These are an information counting model (Tauchen and Pitts 1983), an autoregressive simplification of a diffusion model, here called the ARV model (Taylor 1986), the GARCH(1,1) model (Bollerslev 1986), and the exponential ARCH(1,0) model (Nelson 1991). All four models can explain, first, the high kurtosis and small autocorrelations of daily returns and, second, the statistically significant positive autocorrelations of squared daily returns.

The paper is organized as follows. The terms *stochastic volatility* and *conditional variance* are often used, and they are defined precisely in Section 2. An information counting model is used to show that these two mathematical quantities are in general not functions of each other. Section 3 reviews the continuous-time volatility models that have been used for option pricing. Certain discrete-time simplifications of a popular diffusion model define an AR(1) model for the logarithm of volatility and hence an *autoregressive random variance* (ARV) model for returns. A summary of the model's moments is provided. Section 4 covers *generalized* and *exponential autoregressive conditional heteroskedasticity* (ARCH) models and their moments. Section 5 summarizes theoretical comparisons for the ARV and ARCH models. There are many statistical similarities but different economic principles relate volatility to information and prices. Section 6 compares ARCH and ARV parameter estimates for daily observations of the DM/$ exchange rate from December 1977 to November 1990. The results from fitting several ARCH models support mean reversion in DM/$ volatility for this period and two subperiods. It is noted in Section 7 that both ARV and ARCH models can contribute constructively to the calculation of fair option prices and conclusions follow in Section 8.

2. DEFINITIONS

Let $P(t)$ denote the price of some asset at time t. Assuming no dividends, define the return from an integer time $t - 1$ to time t by

$$X_t = \ln[P(t)/P(t - 1)].$$

The word *volatility* in finance literature is frequently associated with a quantity σ and prices described by the familiar stochastic differential equation

$$d(\ln P) = \mu \, dt + \sigma \, dW$$

with $W(t)$ a standard Wiener process. When μ and σ are constants, X_t has a normal distribution and

$$X_t = \mu + \sigma U_t \qquad \text{with } U_t \sim \mathrm{N}(0,1). \tag{2.1}$$

Furthermore, the U_t are independent and identically distributed (i.i.d.).

2.1. Stochastic Volatility

Equation (2.1) may be generalized by replacing σ by a positive random variable σ_t to give

$$X_t = \mu + \sigma_t U_t \qquad \text{with } U_t \text{ i.i.d. and } U_t \sim \mathrm{N}(0,1). \tag{2.2}$$

Whenever the returns process $\{X_t\}$ can be represented by (2.2), I will call σ_t the *stochastic volatility* for period t. A normal distribution for $(X_t - \mu)/\sigma_t$ is an essential component of the definition of stochastic volatility adopted here.[2]

The stochastic process $\{\sigma_t\}$ will generate realized volatilities $\{\sigma_t^*\}$, which are in general not observable. For any realization σ_t^*,

$$X_t \mid \sigma_t = \sigma_t^* \sim \mathrm{N}(\mu, \sigma_t^{*2}).$$

A mixture of these conditional normal distributions defines the unconditional distribution of X_t, which has excess kurtosis whenever σ_t has positive variance and is independent of U_t.

2.2. Conditional Variance

Capital letters will often be used to denote random variables, and their outcomes will be represented using the correspondence lowercase letters. Given a set of observed returns, $I_{t-1} = \{x_1, x_2, \ldots, x_{t-1}\}$, the conditional variance for period t is

$$h_t = \mathrm{var}(X_t \mid I_{t-1}).$$

It is important to note that the random variable H_t which generates the observed conditional variance h_t is not in general equal to σ_t^2. We have $H_t \neq \sigma_t^2$ both for the models favored in option pricing literature and for ARCH models having nonnormal conditional distributions.

2.3. An Information Counting Model

The easiest way to use economic theory to motivate changes in volatility assumes that returns are defined by a stochastic number of intraperiod price revisions, as in Clark (1973), Tauchen and Pitts (1983), Harris (1987), Lamoureux and Lastrapes (1990), Gallant, Hsieh, and Tauchen (1991), and Andersen (1993). Suppose there are N_t price revisions during trading day t, each caused by unpredictable information. Let event i on day t change the price logarithm by ω_{it}, with

$$X_t = \mu + \sum_{i=1}^{N_t} \omega_{it}.$$

[2] Andersen (1992) offers volatility definitions which avoid distributional assumptions.

Finally suppose that the ω_{it} are i.i.d. and that they are independent of the random variable N_t, with $\omega_{it} \sim \mathrm{N}(0, \sigma_\omega^2)$. Then

$$\sigma_t^2 = \sigma_\omega^2 N_t \quad \text{and} \quad H_t = \sigma_\omega^2 \mathrm{E}[N_t \mid I_{t-1}]. \tag{2.3}$$

Thus, squared volatility is proportional to the amount of price information for this model. Furthermore, $\sigma_t^2 = H_t$ only if the unrealistic assumption is made that price information up to time $t - 1$ determines the quantity of information during period t.

3. MODELS USED TO PRICE OPTIONS

3.1. Continuous-Time Specifications

Option pricing models have been published for general price processes, in which the asset price $P(t)$ and the volatility $\sigma(t)$ each follow a diffusion process. Several authors give numerical options results for specific processes. Scott (1987, 1991), Wiggins (1987), and Chesney and Scott (1989) suppose the logarithm of the volatility follows the Ornstein–Uhlenbeck (O–U) process, as follows:

$$dP/P = \alpha \, dt + \sigma \, dW_1, \tag{3.1}$$

$$d(\ln \sigma) = \lambda(\xi - \ln \sigma) \, dt + \gamma \, dW_2, \tag{3.2}$$

and

$$dW_1 \, dW_2 = \delta \, dt \tag{3.3}$$

with α, λ, ξ, γ, and δ constant parameters and $(W_1(t), W_2(t))$ a two-dimensional standard Wiener process.

At least four other processes for $\sigma(t)$ have been investigated. These are

$$d\sigma = \lambda(\xi - \sigma) \, dt + \gamma \, dW_2, \tag{3.4}$$

$$d\sigma = \lambda\sigma(\xi - \sigma) \, dt + \gamma\sigma \, dW_2, \tag{3.5}$$

$$d\sigma = \lambda\sigma \, dt + \gamma\sigma \, dW_2, \tag{3.6}$$

and

$$d\sigma = \sigma^{-1}(\xi - \lambda\sigma^2) \, dt + \gamma \, dW_2. \tag{3.7}$$

Stein and Stein (1991) and Heston (1993) give closed-form option valuation formulas, respectively, for (3.4) and (3.7). Hull and White (1987a) give results for (3.5), Hull and White (1987b) and Johnson and Shanno (1987) consider (3.6), and Hull and White (1988) evaluate pricing biases for (3.7).

Bailey and Stulz (1989) and Scott (1992) give general equilibrium results for (3.7), with stochastic α in (3.1) and stochastic interest rates. Both papers show that the choice of volatility model can have nontrivial economic implications for fair option values. Finally, replacing σ in (3.1) by $\nu P^{(\beta/2)-1}$ and letting $\ln(\nu)$ follow the O–U process, as in (3.2), gives the interesting model of Melino and Turnbull (1990), when their price drift parameter α is zero.

It is notable that these models have been motivated by convenience and intuition rather than by studies of observed prices.

3.2. Discrete-Time Specifications

Parameter estimates can be obtained for particular discrete-time approximations of the most popular continuous-time model, defined by (3.1)–(3.3). Wiggins (1987), Chesney and Scott (1989), and Duffie and Singleton (1989) choose to work with

$$\ln(P_t) = \ln(P_{t-1}) + \mu + \sigma_{t-1}U_t \tag{3.8}$$

and

$$\ln(\sigma_t) = \alpha + \phi[\ln(\sigma_{t-1}) - \alpha)] + \theta\eta_t, \tag{3.9}$$

with t now restricted to integer values. Here, μ, α, ϕ, and θ are constants, the pairs (U_t, η_t) are i.i.d. and bivariate normal, and the standard normal variables U_t and η_t have correlation δ. The volatility logarithm follows a stationary, Gaussian, AR(1) process when $-1 < \phi < 1$ and a random walk when $\phi = 1$; also the price logarithm follows a random walk with time-varying conditional innovation variance.

The lagged volatility σ_{t-1} appears in (3.8) because this equation is the Euler approximation to (3.1). Consequently σ_{t-1} (not σ_t) is then the stochastic volatility for period t. However, it can be argued that a more natural simplification of (3.1) is simply

$$\ln(P_t) - \ln(P_{t-1}) = X_t = \mu + \sigma_t U_t. \tag{3.10}$$

The special case defined by (3.9), (3.10), and $\delta = 0$ can be found in Taylor (1986). This special case is almost compatible with the equilibria-counting model presented in Section 2.3, providing we ignore the fact that σ_t and the count N_t are, respectively, continuous and discrete random variables.

The words *autoregressive random variance model* (ARV) are used to describe these models and the additional adjectives *lagged* and *contemporaneous* are used to select (3.8) and (3.10), respectively.

The two innovations (U_t, η_t) in the ARV model prevent observation of the sample values of $\{\sigma_t\}$; we can only observe prices. Furthermore, the variance of a return, conditional upon past returns, is an intractable function for ARV models, unlike the simple functions we will see for ARCH models.

3.3. Kalman Filtering

Scott (1987) and Harvey, Ruiz, and Shephard (1994) have shown that estimates of $\ln(\sigma_t)$ can be obtained by applying the Kalman filter to the logarithms of absolute (mean-adjusted) returns, $L_t = \ln(|X_t - \mu|)$. For the contemporaneous ARV model, from (3.9) and (3.10),

$$L_t = \ln(\sigma_t) + \ln(|U_t|) \tag{3.11}$$

and

$$\ln(\sigma_t) = \alpha(1 - \phi) + \phi\ln(\sigma_{t-1}) + \theta\eta_t. \tag{3.12}$$

The variable $\ln(|U_t|)$ has mean $-0.63518\ldots$ and variance $\pi^2/8$ (Scott 1987), and this variable is uncorrelated with η_t whatever the value of δ (Harvey, Ruiz, and Shephard 1994). It is straightforward to estimate the three parameters α, ϕ, and θ by maximizing the quasi-likelihood function of data $\{l_t = \ln|x_t - \bar{x}|\}$. A refinement of the above Kalman filter specification, which uses the additional information provided by the series $\{\text{sign}(x_t - \bar{x})\}$, can be used to obtain an estimate of δ (Harvey and Shephard 1993).

3.4. Moments

Parameter estimates for the ARV model have often been obtained by matching sample and population moments; consequently, these are described in some detail. The ARV model is stationary if and only if $-1 < \phi < 1$ and then $\ln(\sigma_t)$ has mean α and variance $\beta^2 = \theta^2/(1 - \phi^2)$. *Initially,* consider the stationary, lagged ARV model. The unconditional distribution of the return X_t is a lognormal mixture of normal distributions, as in Clark (1973) and Tauchen and Pitts (1983). As U_t is independent of σ_{t-1}, the mean return is μ, $E[|X_t - \mu|^r] < \infty$ for all $r > 0$ and

$$E[|X_t - \mu|] = (2/\pi)^{1/2}e^{\alpha+\beta^2/2}, \tag{3.13}$$

$$\text{var}(X_t) = e^{2\alpha+2\beta^2}, \tag{3.14}$$

and

$$\text{kurtosis}(X_t) = 3e^{4\beta^2}. \tag{3.15}$$

Clearly the distribution of X_t is leptokurtic and the kurtosis can be arbitrarily large.

The autocorrelations of returns are trivially zero at all positive lags. Using ρ to represent autocorrelation and two subscripts to denote the lag and process, $\rho_{\tau,X} = 0$ for all positive τ. The dependence in the volatility process implies $S_t = (X_t - \mu)^2$ and $L_t = \frac{1}{2}\ln(S_t)$ both define autocorrelated processes. Straightforward algebra shows the autocorrelations of the squares S_t are

$$\rho_{\tau,S} = \frac{[1 + 4\delta^2\theta^2\phi^{2(\tau-1)}]e^{4\beta^2\phi^\tau} - 1}{3e^{4\beta^2} - 1}, \qquad \tau > 0.$$

For small β and/or large ϕ^τ, there is the following approximation when δ is zero (Taylor 1986):

$$\rho_{\tau,S} \approx A(\beta)\phi^\tau, \quad A(\beta) = (e^{4\beta^2} - 1)/(3e^{4\beta^2} - 1). \tag{3.16}$$

As L_t has an ARMA(1,1) representation, it can be proved that

$$\rho_{\tau,L} = B(\beta)\phi^{\tau}, \qquad B(\beta) = 8\beta^2/(8\beta^2 + \pi^2), \quad \tau > 0. \tag{3.17}$$

Often β is approximately estimated to be 0.4 in currency studies and then the constants in (3.16) and (3.17) are $A(\beta) = 0.191$ and $B(\beta) = 0.115$.

The parameter δ is proportional to the covariance of X_{t-1} with both $L_t - \phi L_{t-1}$ (Chesney and Scott 1989) and $|X_t - \mu|$ (Melino and Turnbull 1990).

Second, consider the stationary, contemporaneous ARV model. Equations (3.13)–(3.18) can be used when δ is zero, because the lagged and contemporaneous versions then have identical multivariate distributions. When δ is not zero, the return X_t is μ plus the product of correlated variables σ_t and U_t which ensures more complicated results for the moments than before. For example, the mean return no longer equals μ and the returns are not uncorrelated.

As researchers prefer models that have tractable moments, the lagged version will be assumed for the remainder of this paper whenever $\delta \neq 0$, while $\delta = 0$ will be assumed whenever the contemporaneous version is discussed.

3.5. Parameter Estimates

The stationary ARV model has five parameters: the mean return μ, the mean α, the standard deviation β and autoregressive coefficient ϕ of the process $\{\ln(\sigma_t)\}$, and the correlation δ between the two innovation terms. Estimates of ϕ are particularly interesting as they provide information about the persistence of volatility shocks. This parameter has been estimated using simple moment-matching methods (Taylor 1986), the generalized method of moments (Duffie and Singleton 1989; Melino and Turnbull 1990), and ARMA methodology applied to $l_t = \ln(|x_t - \bar{x}|)$ (Chesney and Scott 1989; Scott 1991), assuming $|\phi| < 1$. The quasi-likelihood maximization method noted in Section 3.3 (Harvey, Ruiz, and Shephard 1994) only assumes $|\phi| \leq 1$. Almost all of the estimates of ϕ in these six studies of daily returns are greater than 0.95 and a substantial proportion of the estimates are greater than 0.99. Bayesian probability intervals for ϕ can be found in Jacquier, Polson, and Rossi (1994).

4. MODELS FOR CONDITIONAL HETEROSKEDASTICITY

4.1. General Specifications

ARCH models describe the conditional variance h_t of the return X_t by a simple function of information known at time $t - 1$. The general structure considered here involves a constant expected return μ, a function f which converts the information in previous returns into conditional variances, an i.i.d. sequence $\{Z_t\}$ and a family of distributions $D(\mu, \sigma^2)$ satisfying the condition $Y \sim D(\mu, \sigma^2)$ implies $(Y - \mu)/\sigma \sim D(0,1)$. The general structure can be summarized by the equations

$$X_t = \mu + H_t^{1/2} Z_t, \qquad H_t = f(X_{t-1}, X_{t-2}, \ldots), \quad Z_t \text{ i.i.d.}, \; Z_t \sim D(0,1). \tag{4.1}$$

The autocovariances of the returns are zero at all positive lags because the conditional mean of X_t does not depend on the information I_{t-1} about previous returns. The comprehensive review papers by Bollerslev, Chou, and Kroner (1992) and Bollerslev, Engle, and Nelson (1993) survey the above class of models and many extensions.[3]

4.2. GARCH Specifications

The GARCH(p, q) model developed by Bollerslev (1986) has conditional variance

$$h_t = a_0 + \sum_{i=1}^{q} a_i(x_{t-i} - \mu)^2 + \sum_{j=1}^{p} b_j h_{t-j}. \tag{4.2}$$

The parameters $a_0, \ldots, a_q, b_1, \ldots, b_p$ are all nonnegative. Most researchers find that the special case $p = q = 1$ is satisfactory.

4.3. Exponential ARCH Specifications

The typical term $a_i(x_{t-i} - \mu)^2$ appearing in (4.2) is a symmetric function of $x_{t-i} - \mu$. However, an asymmetric function will be appropriate if there is an asymmetric relationship between volatility and price news. Campbell and Hentschel (1992) refine this idea by developing an economic model. Following Nelson (1991), asymmetry can be modeled by using the following function, $g(z_t)$, of the standardized returns, $z_t = (x_t - \mu)/h_t^{1/2}$, as the residuals of an ARMA(p,q) process for the logarithm of the conditional variance:

$$g(z_t) = \omega z_t + \gamma\{|z_t| - \mathrm{E}[|Z_t|]\}. \tag{4.3}$$

An important example is the AR(1) specification

$$\ln(h_t) - \alpha_t = \Delta[\ln(h_{t-1}) - \alpha_{t-1}] + g(z_{t-1}), \tag{4.4}$$

which will be referred to as the exponential ARCH(1,0) model. The α_t are constants which are typically used to model weekend and holiday effects.

4.4. The Conditional Distribution

When ARCH model parameters are replaced by maximum likelihood estimates the standardized variables z_t usually display excess kurtosis which is a theoretical prediction of the information counting model (Gallant, Hsieh, and Tauchen 1991). Thus, nonnormal conditional distributions are required.

One plausible conditional distribution is a scaled t-distribution. The density function $f(z)$ for D(0,1) depends on the degrees-of-freedom (d.o.f.) parameter ν, as follows:

$$f(z) = \pi^{-1/2}(\nu - 2)^{-1/2}\Gamma\left[\frac{\nu+1}{2}\right]\Gamma\left[\frac{\nu}{2}\right]^{-1}\left\{1 + \frac{z^2}{\nu - 2}\right\}^{-(\nu+1)/2} \quad \text{for } \nu > 2. \tag{4.5}$$

[3]These include multivariate models and models that make use of further information, for example, interest rates related to the stock market's settlement process (Baillie and DeGennaro 1989), trading volume (Lamoureux and Lastrapes 1990), option implied volatilities (Day and Lewis 1992), and FED policy change dummy variables (Lastrapes 1989).

Another plausible conditional distribution is the generalized error distribution (GED) described by Box and Tiao (1973). The density function then depends on a single tail-thickness parameter ν:

$$f(z) = \frac{\nu \exp[-\frac{1}{2}|z/\lambda|^{\nu}]}{\lambda 2^{1+1/\nu}\Gamma[1/\nu]}, \qquad \nu > 0, \tag{4.6}$$

with

$$\lambda = \left\{\frac{2^{-2/\nu}\Gamma[1/\nu]}{\Gamma[3/\nu]}\right\}^{1/2}.$$

When $\nu = 2$, $Z_t \sim \mathrm{N}(0,1)$, while for $\nu < 2$ the distribution has thicker tails than the normal. In particular, $\nu = 1$ gives the double exponential distribution.[4]

4.5. Stochastic Volatility

A t-distribution is a mixture of normal distributions having different variances. From Praetz (1972) and Bollerslev (1987) there is a mixing variable M_t, distributed as inverted gamma, for which

$$X_t - \mu = H_t^{1/2}Z_t = H_t^{1/2}M_t^{1/2}U_t \qquad \text{with } \mathrm{E}[M_t] = 1 \quad \text{and} \quad U_t \sim \mathrm{N}(0,1). \tag{4.7}$$

Then

$$X_t \mid I_{t-1} \sim \mathrm{D}(\mu,h_t) \qquad \text{and} \qquad X_t \mid I_{t-1}, m_t \sim \mathrm{N}(\mu,m_th_t). \tag{4.8}$$

Equations (4.7) and (4.8) also apply for a conditional GED when $1 \le \nu < 2$ (Hsu 1982). Unfortunately, the distribution of the mixing variable M_t, which depends on the tail-thickness parameter ν, appears to be very complicated for the GED; some diagrams are given by Hsu (1980).

Comparing (4.7) with (2.3) shows the stochastic volatility for an ARCH model with conditional t or GED distributions is

$$\sigma_t = (M_tH_t)^{1/2}. \tag{4.9}$$

The distinction between H_t and σ_t^2 can be ascribed to uncertainty about the amount of news revealed during period t. This effect is measured by M_t. When $m_t > 1$, price volatility during period t exceeds the value h_t expected at time $t - 1$, and vice versa. Neither m_t nor the realized value of σ_t can be observed; only h_t is observable.

4.6. Likelihoods

The considerable interest shown in ARCH models is due to their descriptive successes and the availability of maximum likelihood estimates. The likelihood L of returns $\{x_1,$

[4]The expectation of $|Z_t|$ equals $\sqrt{2/\pi}$ when D( ) is normal, $\Gamma[2/\nu]/\{\Gamma[1/\nu]\Gamma[3/\nu]\}^{1/2}$ when D( ) is the GED, and $2(\nu - 2)^{1/2}\Gamma(\nu/2 + 1)/\{\pi^{1/2}(\nu - 1)\Gamma(\nu/2)\}$ when D( ) is a scaled-t distribution.

$x_2, \ldots, x_n\}$, for initial values I_0 and a model parameterized by a vector θ, is the product of conditional densities defined by conditional distributions $D(\mu, h_t)$, with the h_t being functions of the data, I_0 and θ. When $f(z)$ is the density function of D(0,1), which can depend on a subset of θ, the log-likelihood is

$$\ln L(\theta) = \sum_{t=1}^{n} -\frac{1}{2}\ln[h_t(\theta)] + \ln[f(z_t \mid \theta)]. \tag{4.10}$$

Likelihood calculations can be used to compare model specifications and distributions and to test if an integrated model is appropriate. Examples are given in Section 6. Quasi-ML estimates are given by assuming D( ) is normal. These estimates are asymptotically consistent for any D( ) and robust standard errors can be calculated, providing certain regularity conditions are satisfied (Bollerslev and Wooldridge 1992).

4.7. Parameter Estimates

Bollerslev, Chou, and Kroner (1992) provide a review of parameter estimates. Some innovative stock studies are Akgiray (1989), Baillie and DeGennaro (1990), Schwert and Seguin (1990), all of which involve a GARCH specification, and Nelson (1991) which uses exponential ARCH. Two important currency studies are Engle and Bollerslev (1986) and Hsieh (1989). The half-lives of volatility shocks implied by ARCH parameter estimates vary considerably for the studies mentioned, from about two months to several years. It may be expected for daily returns that $a + b$ and Δ are greater than 0.95, respectively, for GARCH (1,1) and exponential ARCH(1,0) estimates.

4.8. Autocorrelations

These are helpful for illustrating similarities with the ARV models of Section 3. First, consider the GARCH(1,1) model. Let $a = a_1$, $b = b_1$ with $a, b > 0$. When $a + b < 1$, the model is covariance stationary, $E[X_t] = \mu$ and $\operatorname{var}(X_t) = a_0/(1 - a - b)$. The X_t have finite kurtosis if and only if

$$k_Z = E[Z_t^4] < \infty \qquad \text{and} \qquad (a + b)^2 + a^2[k_Z - 1] < 1.$$

Providing the X_t have finite kurtosis, the squares have autocorrelations

$$\rho_{\tau,S} = C(a, b)(a + b)^\tau, \qquad \tau > 0,$$

with

$$C(a, b) = \frac{a(1 - ab - b^2)}{(a + b)(1 - 2ab - b^2)}.$$

Second, consider the exponential ARCH(1,0) model. This model is strictly stationary if $|\Delta| < 1$ and the α_t in (4.4) are the same for all t; all the moments of returns are then finite if $D(0,1) \sim$ GED and $\nu > 1$ (Nelson 1991). Formulas for the moments of the squares are extremely complicated. The autocorrelations of

$$L_t = \ln(|X_t - \mu|) = 0.5\ln(H_t) + \ln(|Z_t|)$$

follow from the linear model for $\ln(H_t)$. When D( ) is symmetric,

$$\rho_{\tau,L} = F(\gamma, \Delta, f)\Delta^{\tau}, \qquad \tau > 0,$$

with $F(\gamma, \Delta, f)$ a constant which depends on the density function $f(z)$ of D(0,1), providing these autocorrelations exist; F is positive when γ and Δ are both positive.

5. COMPARISONS

5.1. Similarities

The ARV models of Section 3 have many similarities with particular ARCH models. These similarities are most obvious for the contemporaneous ARV and exponential ARCH(1,0) models. The two models both state that $X_t - \mu$ is the product of an i.i.d., standardized process (either U_t in (3.10) or Z_t in (4.1)) and a second process whose logarithm is linear and first-order autoregressive (either σ_t defined by (3.9) or $H_t^{1/2}$ defined by (4.4)). The only obvious difference between the two models is that the product involves processes which are independent of each other for the contemporaneous ARV model but not for the exponential ARCH model.

Stationary ARV, GARCH(1,1) and exponential ARCH(1,0) models can all explain the following "stylized facts" for samples of daily returns: (1) they have excess kurtosis compared with the normal distribution, (2) they have small autocorrelations, and (3) the autocorrelations of squared returns are (i) positive, (ii) statistically significant, (iii) generally smaller as the time-lag increases, and (iv) noticeably larger than the respective autocorrelations of returns for several time lags. The population autocorrelations are very similar for all the models. All models have $\rho_{\tau,X} = 0$ for all positive τ, while the squares $S_t = (X_t - \mu)^2$ have

$$\rho_{\tau,S} \approx A(\beta)\phi^{\tau} \qquad \text{for the ARV model,}$$

and

$$\rho_{\tau,S} = C(a, b)(a + b)^{\tau} \qquad \text{for GARCH(1,1),} \quad \text{when } \tau > 0,$$

with $A, C > 0$. Also the variables $L_t = \frac{1}{2}\ln(S_t)$ have

$$\rho_{\tau,L} = B(\beta)\phi^{\tau} \qquad \text{for the ARV model,}$$

and

$$\rho_{\tau,L} = F(\gamma,\Delta,f)\Delta^{\tau} \qquad \text{for exp. ARCH(1,0),} \qquad \tau > 0,$$

with $B, F > 0$ (assuming $\gamma,\Delta > 0$ and symmetric conditional distributions). The preceding autocorrelations are those of ARMA(1,1) models. It follows that any application of standard identification methods to the autocorrelations, partial autocorrelations, and/or spectral densities of the squares or their logarithms cannot provide information about the relative validity of the ARV and ARCH models. This conclusion is a natural consequence of a theorem proved by Nelson (1990), which shows that a particular sequence of exponential

ARCH(1,0) models converges to the bivariate diffusion process for price and volatility which motivates the ARV model ((3.1)–(3.3)).[5]

5.2. Discussion

Although the ARV and exponential ARCH(1,0) models are approximations to the same continuous-time process, it does not follow that the approximations are of equal value for describing prices observed at discrete moments in time. It is difficult to decide which model provides the best discrete-time approximation, either theoretically or by studying data. At present it is unknown if this uncertainty matters.

A probabilistic distinction between the models makes use of the idea of a *reversible* stochastic process. Kelly (1979, p. 5) writes of such a process: "... when the direction of time is reversed the behaviour of the process remains the same...."[6] A Gaussian AR(1) process is reversible. Consequently the contemporaneous ARV model *is* reversible, because it is the product of stochastically independent reversible processes. However, the exponential ARCH(1,0) model is *not* reversible. This model has characteristics which are different when it is studied backwards through time.[7]

5.3. Implications for the Information Counting Model

A first impression is that prices cause volatility in ARCH models but not in the contemporaneous ARV model. As volatility is a function of information for the equilibria-counting model of Section 2.3, it is tempting to conclude that prices cause the number of future information flows for ARCH models but not for the ARV model.

To be more precise, first consider the volatility residuals for the ARV model:

$$\varepsilon_t = \ln(\sigma_t) - \phi\ln(\sigma_{t-1}) - (1 - \phi)\alpha,$$

with ϕ possibly equal to 1. Then ε_t is independent of the returns history I_{t-1}. This implies that prices cannot "cause" volatility. Second, for the exponential ARCH(1,0) model consider

$$\varepsilon_t^* = \ln(\sigma_t) - \Delta\ln(\sigma_{t-1}) - 0.5(1 - \Delta)\alpha_E,$$

with Δ possibly equal to 1. Using (4.4) and (4.9) it then follows that

$$2\varepsilon_t^* = g(Z_{t-1}) + \ln(M_t) - \Delta\ln(M_{t-1}).$$

Consequently, the adjustment ε_t^* depends on the information set I_{t-1}. Thus prices partially "cause" volatility for this model, through the term Z_{t-1} which is a function of the entire history of returns. For conditional normal distributions, $M_t = 1$ so that ε_t^* is then a deterministic function of past prices.

[5] The theorem suggests methods for estimating the parameters of the diffusion model. For example, the correlation δ between the price and volatility differentials in (3.3) can be estimated by fitting (4.3) and (4.4), with conditional normal distributions, because $\delta \approx \operatorname{sign}(\omega)[1 + (\gamma/\omega)^2(1 - 2/\pi)]^{-1/2}$.

[6] A stochastic process is reversible if and only if the likelihood function has $L(x_{t+1}, x_{t+2}, \ldots, x_{t+n}) = L(x_{t+n}, x_{t+n-1}, \ldots, x_{t+1})$ for all $n, t, x_{t+1}, \ldots, x_{t+n}$.

[7] Simulations have shown there are several such characteristics when the conditional distribution is normal but the results to date are not constructive when the conditional distribution is double exponential (the GED with $\nu = 1$).

Suppose we now accept that the intraperiod price revision model of Section 2.3 is appropriate. Then the information count N_t is proportional to σ_t^2. It follows that the number of information items absorbed by the market during a period is a function of the previous period's number and a count residual term which is independent of past prices for the ARV model but dependent on past prices for the exponential ARCH model (but only conceptually as N_t is discrete while σ_t^2 is a continuous random variable). To conclude:

(i) Prices change when information becomes available to a market,
(ii) The contemporaneous ARV model states that, given N_{t-1}, how much the price has changed in period $t - 1$ is irrelevant for determining future volatility and the number of future information items (N_t only depends on constants, N_{t-1} and ε_t),
(iii) In contrast, ARCH models imply that the magnitude of price changes partially determines future volatility and the number of future information items (N_t depends on N_{t-1}, I_{t-1}, M_t, and M_{t-1}).

6. RESULTS FOR THE DM/$ EXCHANGE RATE

6.1. Data

Model parameters have been estimated for a selection of ARCH models and the contemporaneous ARV model, permitting a comparison of the results from different models and estimation methods. The results are for daily DM/$ returns covering the 13 years from December 1977 to November 1990. These changes in the price logarithm are calculated from the closing prices of June and December futures contracts traded in Chicago. Six months of prices are used for each contract, commencing with prices from December 1977 to May 1978 for the June 1978 contract.[8] Models have been estimated for the complete time series of futures returns (3283 observations) and also for two subsets (Dec. 1977–Dec. 1983, 1533 observations and Jan. 1984–Nov. 1990, 1750 observations).

6.2. ARCH Methodology

An appraisal of calendar dummy variables showed that seven are required. Variables for Mondays and holidays are included in the conditional variance equations because returns measured over more than 24 hours have higher variances than 24-hour returns. Five additional variables define the conditional means, one for each day of the week (cf. McFarland, Pettit, and Sung 1982).

The ARCH models are fitted to returns minus their conditional means. A nonseasonal conditional variance h_t^* is defined to be the conditional variance h_t divided by any relevant dummy variable, thus:

$$
\begin{aligned}
h_t/h_t^* &= 1 && \text{if close } t \text{ is 24 hours after close } t-1,\\
&= M && \text{if } t \text{ falls on a Monday and } t-1 \text{ on a Friday,}\\
&= V && \text{if a holiday (vacation) occurs between close } t \text{ and close } t-1.
\end{aligned}
$$

[8] Every return is calculated using two prices for the same contract. On a few occasions the futures prices reported by the market are not transaction prices because of the limit rules operated until May 1985. Consequently, 28 futures prices have been replaced by equivalent forward prices; 17 of these replacements are needed in 1978.

Nonseasonal ARCH models for h_t^* combined with the above definition of h_t/h_t^* define appropriate seasonal models for h_t.

Only one initial value is required in the calculations for most of the models fitted here, namely h_1, or equivalently h_1^*. I have used all the observations to define h_1^* to be the appropriate estimate of the unconditional, nonseasonal variance.

6.3. ARCH Results

Comparisons are first made between the GARCH(1,1) model and the symmetric exponential ARCH(1,0) model and between the three most popular conditional distributions: the normal, the scaled-t, and the GED. Table 6.1 presents the results for the complete series from fitting the six combinations of ARCH specification and conditional distribution.

The results show that the GED is superior to the t-distribution for describing the conditional distributions of DM/$ returns during the period studied. The maximum log-likelihood is 10.51 more for GARCH-GED than for GARCH-t and it is 9.34 more for exp.-ARCH-GED than for exp.-ARCH-t. The differences range from 3.91 to 4.40 for the two subperiods.[9] Conditional normal distributions are unacceptable. The exponential ARCH(1,0) model then has a maximum log-likelihood which is 73.27 less than the figure for a conditional GED specification. As the normal distribution is the GED with $\nu = 2$, doubling the difference to 146.54 and comparing with χ_1^2 makes it clear that a nonnormal conditional distribution is essential.

Comparing the maximum log-likelihoods for GARCH with exponential ARCH reveals that the latter fits slightly better with log-likelihood differences of 5.01 and 6.18, respectively, for the GED and t-distribution specifications. The differences range from 0.77 to 1.76 for the subperiods. The estimates of M and V are approximately 1.4 and 1.7, respectively, and are greater than the comparable stock estimates in French and Roll (1986) and Nelson (1991). The choice of initial value of h_1 does not matter much. Letting h_1 be an additional parameter increases the maximum log-likelihood by only 0.33 for GARCH-t and 0.11 for exp.-ARCH-GED.

6.4 Further Exponential-ARCH-GED Results

The best model from Table 6.1, namely the symmetric exponential ARCH(1,0)-GED model, is now taken to be a benchmark against which other models can be compared.[10] The adjusted log-likelihood for an alternative model, denoted AL, is defined to be the maximum log-likelihood for that model minus the comparable value for the benchmark model.

First, consider whether an integrated model is credible. Letting $\Delta = 1$ in (4.4) gives a driftless random walk for $\ln(H_t)$. Fitting this model (with $\omega = 0$) and allowing h_1 to be a parameter gives $AL = -19.11$. Including a further parameter for any drift in the random walk for $\ln(H_t)$ leaves AL unchanged. The likelihood ratio test using $-2AL$ as the test statistic very probably has size greater than the significance level and is therefore unreli-

[9] It is not possible to immediately deduce that the GED is preferable, because the likelihood functions are not nested. Estimation of an exp.-ARCH model in which Z_t has density $pf_1(z) + (1 - p)f_2(z)$, with f_1 the t density and f_2 the GED density, provides no evidence to reject $p = 0$ but obviously rejects $p = 1$. My differences conflict with the claim made by Baillie and Bollerslev (1989) that the t-distribution gives substantially higher maximum log-likelihoods than the GED for their currency data. However, the term $-\frac{1}{2}\ln(\pi)$ is missing from their log-likelihood function for the t-distribution.

[10] The standardized residuals for this model are compatible with the model assumptions. The average values of z_t, $z_t^2 - 1$, z_t^3, $z_t z_{t+\tau}$, and $(z_t^2 - 1)(z_{t+\tau}^2 - 1)$, $1 \leq \tau \leq 10$, are all close to zero as required.

Parameter Estimates for GARCH(1,1) and Exponential ARCH(1,0) Models

Parameter	GARCH Estimates			Exponential ARCH Estimates		
	Normal	t	GED	Normal	t	GED
a	0.0990	0.0892	0.0935			
γ				0.1983	0.1805	0.1885
(s.e.)[a]	(0.0117)	(0.0137)	(0.0141)	(0.0199)	(0.0231)	(0.0239)
[r.s.e.][b]	[0.0165]			[0.0295]		
$a + b$	0.9702	0.9790	0.9737			
Δ				0.9607	0.9702	0.9658
(s.e.)	(0.0083)	(0.0092)	(0.0098)	(0.0076)	(0.0080)	(0.0087)
[r.s.e.]	[0.0111]			[0.0117]		
$10^5 a_0/(1 - a - b)$	5.270	5.707	5.283			
α_E				−9.827	− 10.013	− 10.071
(s.e.)	(0.703)	(1.384)	(0.954)	(0.086)	(0.096)	(0.095)
[r.s.e.]	[0.860]			[0.111]		
$1/\nu$		0.155			0.158	
ν	2		1.321	2		1.317
(s.e.)	(0)	(0.018)	0.047)	(0)	(0.018)	(0.046)
M	1.407	1.430	1.423	1.389	1.415	1.406
(s.e.)	(0.085)	(0.105)	(0.106)	(0.086)	(0.106)	(0.107)
[r.s.e]	[0.109]			[0.109]		
V	1.519	1.731	1.643	1.550	1.792	1.695
(s.e.)	(0.198)	(0.283)	(0.271)	(0.206)	(0.294)	(0.283)
[r.s.e.]	[0.250]			[0.237]		
max.ln(L)	11738.49	11798.46	11808.97	11740.71	11804.64	11813.98

Maximum likelihood estimates and the maxima of the log-likelihood function for ARCH models fitted to 3283 daily DM/$ returns (December 1977 to November 1990). The ARCH models contain seasonal dummy variables M and V for Mondays and vacations. The conditional distribution is either normal, Student's t, or generalized exponential. Five day-of-the-week conditional mean terms are included in the models. The estimates of these parameters are not given here.

Models for the conditional variance h_t: $h_t = h_t^*$, Mh_t^*, or Vh_t^*; $x_t^* = x_t - \mathrm{E}[X_t \mid I_{t-1}]$;

GARCH(1,1) $\quad h_t^* = a_0 + a(x_{t-1}^*)^2 + bh_{t-1}^*$,

exp. ARCH(1,0) $\quad \ln(h_t^*) = \alpha_E + \Delta[\ln(h_{t-1}^*) - \alpha_E] + \gamma(|x_{t-1}^*| / \sqrt{h_{t-1}} - \text{cst.})$.

[a] s.e. refers to the standard error. These are calculated using the Hessian and numerical second derivatives.

able, but a robust Wald test for a unit root in the volatility process appears to be reliable (Lumsdaine 1991a,b). The test statistic can be calculated from quasi-ML estimates and formulas given by Bollerslev and Wooldridge (1992) and Engle and Gonzalez-Rivera (1991). From Table 6.1, the robust Wald statistic is $t = (\hat{\Delta} - 1)/\text{s.e.}\ (\hat{\Delta}) = (0.9607\text{-}1)/0.0117 = -3.36$, which is to be compared with N(0,1).[11] It is concluded that DM volatility was mean-reverting during the period from 1978 to 1990. The data support this conclusion for both subperiods using a 5% significance level ($t = -2.96$, 1978–83; $t = -2.34$, 1984–90).

Second, consider alternative specifications for the volatility residual $g(z_{t-1})$. The benchmark model has $\omega = 0$ in the equation

$$g(z_{t-1}) = \omega z_{t-1} + \gamma(|z_{t-1}| - E[|Z_{t-1}|]). \tag{6.1}$$

Specification (6.1) and

$$g(z_{t-1}) = \omega z_{t-1} + \gamma(z_{t-1}^2 - 1), \tag{6.2}$$

respectively, have $AL = 0.55$ and $AL = -10.51$. Thus the kinked function (6.1) helps to describe DM returns better than the smooth function (6.2). Considering (6.1) further, the estimated values of ω and γ are 0.0119 and 0.1883, respectively, and the small value of AL clearly supports the hypothesis $\omega = 0$. The estimated ratio ω/γ is only 0.06 compared with the U.S. stock values of -0.76 in Nelson (1991) and -0.39 in Brock, Lakonishok, and LeBaron (1992). The result here is not surprising: there are plausible theories for a negative ω in stock models but none for a nonzero ω in currency models.

Third, consideration of exponential ARCH(p,q) models with $(p,q) \neq (1,0)$ shows that they have nothing extra to offer. For example, $(p,q) = (2,1)$ has AL equal to only 0.55.

Table 6.2 presents subperiod estimates for the benchmark model. The subperiod estimates and their standard errors suggest there may have been significant changes in the

(i) Magnitude of volatility shocks ($\hat{\gamma}$ decreases from 0.27 to 0.14, t-ratio -2.58)
(ii) Long-run level of volatility ($\hat{\alpha}_E$ increases from -10.34 to -9.87, t-ratio 2.80)
(iii) Weekend volatility effect ($\hat{M}$ decreases from 1.74 to 1.18, t-ratio -2.45)
(iv) Thursday conditional mean term (t-ratio 2.53)

Quasi-ML estimates and their robust standard errors give similar t-ratios for the estimated parameter changes: -1.96 for $\hat{\gamma}$, 2.13 for $\hat{\alpha}_E$, -2.49 for $\hat{M}$, and 2.70 for the Thursday mean.

The point estimates of α_E, M, and V imply the median annualized volatility is 10% for the first subperiod and 12% for the second period. Fitting models in which α_E depends linearly on the time t, so $\alpha_E = \alpha_0 + \alpha_1 t$, gives $\hat{\alpha}_1 > 0$ for the complete dataset but $\hat{\alpha}_1 < 0$ for each subperiod. There is therefore little evidence for an upward drift in DM/$ volatility.

6.5 ARV Estimates

Seasonal effects are included in the ARV model by writing $\ln(\sigma_t) - \ln(\sigma_t^*) = 0$, $0.5\ln(M)$, or $0.5\ln(V)$, as appropriate, and then supposing that $\{\ln(\sigma_t^*)\}$ is a Gaussian,

[11] $t = -2.68$ for the GARCH(1,1) specification.

TABLE 6.2
Exponential ARCH(1,0) Estimates for Subperiods

Parameter	Complete Dataset (Dec. 1977–Nov. 1990; 3283 obs.)	First Subperiod (Dec. 1977–Dec. 1983; 1533 obs.)	Second Subperiod (Jan. 1984–Nov. 1990; 1750 obs.)
γ	0.1885	0.2659	0.1368
(s.e.)[a]	(0.0239)	(0.0424)	(0.0267)
Δ	0.9658	0.9478	0.9628
(s.e.)	(0.0087)	(0.0167)	(0.0128)
α_E	−10.071	−10.337	−9.871
(s.e.)	(0.095)	(0.132)	(0.101)
ν	1.317	1.304	1.357
(s.e.)	(0.046)	(0.066)	(0.066)
M	1.406	1.735	1.180
(s.e.)	(0.107)	(0.191)	(0.122)
V	1.695	2.054	1.414
(s.e.)	(0.283)	(0.497)	(0.324)
Mean return $\times 10^4$			
Monday	−6.26	−7.94	−4.86
(s.e.)	(2.63)	(3.44)	(3.68)
Tuesday	−2.20	−4.18	0.23
(s.e.)	(2.12)	(2.73)	(3.03)
Wednesday	4.43	5.16	2.76
(s.e.)	(2.18)	(2.19)	(3.46)
Thursday	−4.57	−9.37	2.20
(s.e.)	(2.21)	(2.99)	(3.47)
Friday	−4.26	−6.00	−2.67
(s.e.)	(2.35)	(3.06)	(3.10)
max. $\ln(L)$	11813.98	5683.58	6146.94

Maximum likelihood estimates and the maxima of the log-likelihood function for exponential ARCH(1,0) models fitted to daily DM/$ returns for various periods. The ARCH models contain seasonal dummy variables M and V for Mondays and vacations. Five day-of-the-week conditional mean terms are included in the models. The conditional distribution is generalized exponential with tail-thickness parameter ν.

Model for the conditional variance h_t: $h_t = h_t^*$, Mh_t^*, or Vh_t^*; $x_t^* = x_t - E[X_t|I_{t-1}]$,

$$\ln(h_t^*) = \alpha_E + \Delta[\ln(h_{t-1}^*) - \alpha_E] + \gamma(|x_{t-1}^*|/\sqrt{h_{t-1}} - \text{cst.}).$$

[a] s.e. refers to the standard error. These are calculated using the Hessian and numerical second derivatives.

AR(1) process having mean α, standard deviation β, and autoregressive parameter ϕ. Many methods for estimating α, β, and ϕ were noted in Section 3.5. Some of these methods are used here. Other methods, including GMM, are the subject of continuing research.

The estimation methods make use of adjusted returns x_t^*, defined by $x_t^* = x_t - \bar{x}$, $(x_t - \bar{x})/M^{1/2}$, or $(x_t - \bar{x})/V^{1/2}$, as appropriate, with $M = 1.406$ and $V = 1.695$ from Table 6.1. Matching the mean absolute deviation and the variance of $\{x_t^*\}$, using (3.13) and (3.14), gives $\hat{\alpha} = -5.153$ and $\hat{\beta} = 0.415$. Alternatively, β can be estimated by matching the kurtosis using (3.15) and, on this occasion, a similar estimate is obtained: $\hat{\beta} = 0.412$.

The autoregressive parameter ϕ has been estimated in three ways. Fitting the function $K\phi^{\tau}$ to the first 50 autocorrelations of the absolute adjusted returns, $|x_t^*|$, by ordinary least squares, gives $\hat{K} = 0.114$ and $\hat{\phi} = 0.969$. Second, fitting a linear ARMA(1,1) model to the variables $l_t = \ln(|x_t^*|)$ gives $\hat{\phi} = 0.972$ with an estimated standard error of 0.026. The standard error of $\hat{\phi}$ is more than twice as large as the comparable ARCH standard errors and a simple test of the unit-root hypothesis $\phi = 1$ would accept the possibility that the volatility was not mean reverting. A Type II error would then occur according to the ARCH test described in Section 6.4. Third, maximizing the quasi-likelihood function of the series $\{l_t\}$, using the Kalman filter specification given by (3.11)–(3.12), gives $\hat{\phi} = 0.938$ and also $\hat{\alpha} = -5.193$ and $\hat{\beta} = 0.404$. Subperiod estimates show an increase in $\hat{\alpha}$ and a decrease in $\hat{\beta}$ from the first to the second subperiod, approximately comparable to the changes in the estimates of the exponential ARCH parameters α_E and γ.

6.6. Comparisons of Volatility Persistence Estimates

The sum $a + b$, for a GARCH(1,1) model, is a measure of the persistence of volatility shocks, likewise Δ for the exponential ARCH(1,0) model and ϕ for the ARV model. Table 6.3 presents seven estimates of these persistence measures for the complete dataset and seven estimates for each of the two subperiods. It can be seen that six of the full-period estimates are similar but the Kalman filter estimate is somewhat smaller, although all the estimates are similar for the subperiods. In six of the rows in Table 6.3 both the subperiod estimates are slightly less than the full-period estimate, which may reflect one or more changes in the unconditional variance as noted earlier.

An estimate of $a + b$, Δ, or ϕ equal to 0.9669 (the average figure in Table 6.3 for all 13 years) implies a half-life h of 30 calendar days by solving the equation $(0.9669)^{252h/365} = 0.5$. For an estimate equal to 0.9610 (the subperiod average figure), the half-life is 25 calendar days.

7. OPTION PRICING IMPLICATIONS

The ARV and exponential ARCH(1,0) models provide fair option prices which are almost identical when the model parameters are matched appropriately. Nevertheless, both the ARV and ARCH models are constructive when valuing options.

Numerical valuation methods are much quicker for the ARV model because it typically requires integration of the Black–Scholes function multiplied by the probability density of the average variance during the life of the option, if it is assumed that (i) volatility risk is not priced and (ii) price and volatility innovations are uncorrelated (Hull and White 1987a). Integration over a variance distribution is much faster than the ARCH alternative which requires integration over a price distribution with respect to a risk-neutral measure, if similar assumptions about risk and the innovations are made (Duan 1993; Engle and Mustafa 1992). These integrals have been evaluated using Monte Carlo methods enhanced by the antithetic variable and control-variate methods. Approximately 2000 option prices can be calculated from the ARV model in the time taken to calculate one option price from the exponential ARCH(1,0) model to the same accuracy, when the volatility parameters are similar to the DM/$ estimates.

TABLE 6.3
Persistence Estimates for DM/$ Volatility

Model[a]	Parameter	Complete Dataset (Dec. 1977–Nov. 1990; 3283 obs.) Estimate	First Subperiod (Dec. 1977–Dec. 1983; 1533 obs.) Estimate	Second Subperiod (Jan. 1984–Nov. 1990; 1750 obs.) Estimate
ARV	ϕ^b	0.9384	0.9574	0.9519
Exp-ARCH-GED	Δ	0.9658	0.9478	0.9628
ARV	ϕ^c	0.9688	0.9667	0.9583
Exp-ARCH-t	Δ	0.9702	0.9558	0.9656
ARV	ϕ^d	0.9719	0.9646	0.9666
GARCH-GED	$a + b$	0.9737	0.9645	0.9581
GARCH-t	$a + b$	0.9790	0.9728	0.9612
Average		0.9669	0.9613	0.9606

Estimates of persistence parameters for ARCH and ARV models fitted to daily DM/$ returns. The ARCH estimates are maximum likelihood estimates. The order of the rows in the table is determined by the magnitudes of the persistence estimates for the complete dataset.

[a]Exp-ARCH refers to exponential ARCH(1,0), GARCH to GARCH(1,1), GED to conditional generalized exponential distributions, t to conditional Student's t-distributions, and ARV to the contemporaneous autoregressive random variance model.

[b]Estimates obtained by the Kalman filter method of Harvey, Ruiz, and Shephard (1994).

[c]Estimates obtained by matching autocorrelations, as in Taylor (1986).

[d]Estimates obtained by the ARMA fitting method of Chesney and Scott (1989).

Model estimation and validation is simpler when ARCH models are used. These models permit robust tests for a unit root in the volatility process and tests of the assumption that price and volatility residuals are uncorrelated, as illustrated here in Section 6.4. Comparisons of different ARCH specifications provide guidance about a suitable volatility process. The results given here show exponential ARCH(1,0) describes the DM/$ returns best within the ARCH family and hence the ARV model is a credible choice when the distribution of the average variance is required. An initial variance is needed when the average variance is simulated. ARCH equations provide an appropriate value for this initial variance.

8. CONCLUSIONS

The theoretical results summarized in Section 5, theorems proved by Nelson (1990), and the empirical results reported here and elsewhere all support an important conclusion: a discrete-time approximation to a stochastic volatility model used to price options in theoretical studies (here called the ARV model) is very similar to the best ARCH models identified in descriptive studies of asset prices. For the exchange rates considered, a methodology which uses both ARV and ARCH models permits the very quick calculation of options values based upon stochastic processes which are empirically credible. A fundamental difference between the two types of models can be identified when volatility is a function of the quantity of information incorporated into prices: variation in this quantity exclusively explains changes in volatility for the ARV model while past prices are the primary determinant of volatility changes for ARCH models.

REFERENCES

AKGIRAY, V. (1989): "Conditional Heteroscedasticity in Time Series of Stock Returns: Evidence and Forecasts," *J. Business*, 62, 55–80.

ANDERSEN, T. G. (1992): "Volatility," working paper, Kellogg Graduate School of Management, Northwestern University.

ANDERSEN, T. G. (1993): "Return Volatility and Trading Volume: An Information Flow Interpretation of Stochastic Volatility," working paper, Kellogg Graduate School of Management, Northwestern University.

BAILEY, W., and R. M. STULZ (1989): "The Pricing of Stock Index Options in a General Equilibrium Model," *J. Financial Quant. Anal.*, 24, 1–12.

BAILLIE, R. T., and T. BOLLERSLEV (1989): "The Message in Daily Exchange Rates: A Conditional-Variance Tale," *J. Business Econ. Statis.*, 7, 297–305.

BAILLIE, R. T., and R. P. DEGENNARO (1989): "The Impact of Delivery Terms on Stock Return Volatility," *J. Financial Services Res.*, 3, 55–76.

BAILLIE, R. T., and R. P. DEGENNARO (1990): "Stock Returns and Volatility," *J. Financial Quant. Anal.*, 25, 203–214.

BOLLERSLEV, T. (1986): "Generalized Autoregressive Conditional Heteroscedasticity," *J. Econometrics*, 31, 307–327.

BOLLERSLEV, T. (1987): "A Conditionally Heteroskedastic Time Series Model for Security Prices and Rates of Return Data," *Rev. Econ. Statis.*, 59, 542–547.

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

BOLLERSLEV, T., R. F. ENGLE, and D. B. NELSON (1993): "ARCH Models," in *The Handbook of Econometrics*, vol. 4, forthcoming.

BOLLERSLEV, T., and J. M. WOOLDRIDGE (1992): "Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time Varying Covariances," *Econometric Rev.*, 11, 143–172.

BOX, G. E. P., and G. C. TIAO (1973): *Bayesian Inference in Statistical Analysis*. Reading, MA: Addison-Wesley.

BROCK, W., J. LAKONISHOK, and B. LEBARON (1992): "Simple Technical Trading Rules and the Stochastic Properties of Stock Returns," *J. Finance*, 47, 1731–1764.

CAMPBELL, J. Y., and L. HENTSCHEL (1992): "No News Is Good News: An Asymmetric Model of Changing Volatility in Stock Returns," *J. Financial Econ.*, 31, 281–318.

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

CLARK, P. K. (1973): "A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices," *Econometrica*, 41, 135–155.

DAY, T. E., and C. M. LEWIS (1992): "Stock Market Volatility and the Information Content of Stock Index Options," *J. Econometrics*, 52, 267–287.

DUAN, J.-C. (1993): "The GARCH Option Pricing Model," *Math. Finance*, 4(2).

DUFFIE, D., and K. J. SINGLETON (1989): "Simulated Moments Estimation of Markov Models of Asset Prices," unpublished manuscript, Graduate School of Business, Stanford University; conference paper, London School of Economics.

ENGLE, R. F. (1982): "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of U.K. Inflation," *Econometrica*, 50, 987–1008.

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

ENGLE, R. F., and G. GONZÁLEZ-RIVERA (1991): "Semiparametric ARCH Models," *J. Business Econ. Statis.*, 9, 345–359.

ENGLE, R. F., and C. MUSTAFA (1992): "Implied ARCH Models from Options Prices," *J. Econometrics*, 52, 289–311.

FRENCH, K. R., and R. ROLL (1986): "Stock Return Variances: The Arrival of Information and the Reaction of Traders," *J. Financial Econ.*, 17, 5–26.

GALLANT, A. R., D. A. HSIEH, and G. E. TAUCHEN (1991): "On Fitting a Recalcitrant Series: The Pound/Dollar Exchange Rate, 1974–83," in *Nonparametric and Semiparametric Methods in Econometrics and Statistics*, eds., W. A. Barnett, J. Powell, and G. E. Tauchen. Cambridge, UK: Cambridge University Press.

HARRIS, L. (1987): "Transaction Data Tests of the Mixture of Distributions Hypothesis," *J. Financial Quant. Anal.*, 22, 127–141.

HARVEY, A. C., E. RUIZ, and N. G. SHEPHARD (1994): "Multivariate Stochastic Variance Models," *Rev. Economic Stud.*, forthcoming.

HARVEY, A. C., and N. G. SHEPHARD (1993): "The Econometrics of Stochastic Volatility," working paper, Department of Statistical and Mathematical Sciences, London School of Economics.

HESTON, S. L. (1993): "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options," *Rev. Financial Stud.*, 6, 327–343.

HSIEH, D. A. (1989): "Modelling Heteroscedasticity in Daily Foreign-Exchange Rates," *J. Business Econ. Statist.*, 7, 307–317.

HSU, D.-A. (1980): "Further Analyses of Position Errors in Navigation," *J. Navigation*, 33, 452–474.

HSU, D.-A. (1982): "A Bayesian Robust Detection of Shift in the Risk Structure of Stock Market Returns," *J. Am. Statis. Assoc.*, 77, 29–39.

HULL, J., and A. WHITE (1987a): "The Pricing of Options on Assets with Stochastic Volatilities," *J. Finance*, 42, 281–300.

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

HULL, J., and A. WHITE (1988): "An Analysis of the Bias in Option Pricing Caused by a Stochastic Volatility," *Adv. Futures Options Res.*, 3, 29–61.

JACQUIER, E., N. G. POLSON, and P. E. ROSSI (1994): "Bayesian Analysis of Stochastic Volatility Models," *J. Business Econ. Statis.*, forthcoming.

JOHNSON, H., and D. SHANNO (1987): "Option Pricing When the Variance Is Changing," *J. Financial Quant. Anal.*, 22, 143–152.

KELLY, F. P. (1979): *Reversibility and Stochastic Networks*. Chichester, UK: Wiley.

LAMOUREUX, C. G., and W. D. LASTRAPES (1990): "Heteroskedasticity in Stock Return Data: Volume versus GARCH Effects," *J. Finance*, 45, 221–229.

LASTRAPES, W. D. (1989): "Weekly Exchange Rate Volatility and US Monetary Policy Regimes: An Application of the ARCH Model," *J. Money Credit Bank.*, 21, 66–77.

LUMSDAINE, R. L. (1991a): "Asymptotic Properties of the Quasi-Maximum Likelihood Estimator in GARCH(1,1) and IGARCH(1,1) Models," working paper, Department of Economics, Princeton University.

LUMSDAINE, R. L. (1991b): "Finite Sample Properties of the Maximum Likelihood Estimator in GARCH(1,1) and IGARCH(1,1) Models: A Monte Carlo Investigation," working paper, Department of Economics, Princeton University.

MCFARLAND, J. W., R. R. PETTIT, and S. K. SUNG (1982): "The Distribution of Foreign Exchange Price Changes: Trading Day Effects and Risk Measurement," *J. Finance*, 37, 693–715.

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

MERVILLE, L. J., and D. R. PIEPTEA (1989): "Stock Price Volatility, Mean-Reverting Diffusion, and Noise," *J. Financial Econ.*, 24, 193–214.

NELSON, D. B. (1990): "ARCH Models as Diffusion Approximations," *J. Econometrics*, 45, 7–38.

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

PRAETZ, P. D. (1972): "The Distribution of Share Price Changes," *J. Business*, 45, 49–55.

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

SCOTT, L. O. (1987): "Option Pricing When the Variance Changes Randomly: Theory, Estimation, and an Application," *J. Financial Quant. Anal.*, 22, 419–438.

SCOTT, L. O. (1991): "Random Variance Option Pricing: Empirical Tests of the Model and Delta-Sigma Hedging," *Adv. Futures Options Res.*, 5, 113–135.

SCOTT, L. O. (1992): "Stock Market Volatility and the Pricing of Index Options: An Analysis of Implied Volatilities and the Volatility Risk Premium in a Model with Stochastic Interest Rates and Volatility," working paper, Department of Finance, University of Georgia.

STEIN, E. M., and J. C. STEIN (1991): "Stock Price Distributions with Stochastic Volatility: An Analytic Approach," *Rev. Financial Stud.*, 4, 727–752.

STEIN, J. C. (1989): "Overreactions in the Options Market," *J. Finance*, 44, 1011–1023.

TAUCHEN, G. E., and M. PITTS (1983): "The Price Variability-Volume Relationship on Speculative Markets," *Econometrica*, 51, 485–505.

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

WIGGINS, J. B. (1987): "Option Values under Stochastic Volatility: Theory and Empirical Estimates," *J. Financial Econ.*, 19, 351–372.