GARCH Modelling of Volatility: An Introduction to Theory and Applications

Theo E. Nijman and Franz C. Palm*
Tilburg University and Limburg University

1. Introduction

The objective of this chapter is to introduce GARCH models of volatility in financial series, discussing the properties of these models and showing how they have been applied in finance and international economics. Rather than presenting a complete survey of this rapidly expanding literature, as provided by Bollerslev, Chou and Kroner (1992), we discuss a limited number of applications in detail.

This chapter is organized as follows. The relevance of econometric models of volatility for modern finance will be illustrated in section 2. In section 3, the autoregressive conditional heteroskedastic (ARCH) model put forward by Engle (1982), and extensions such as GARCH, ARCH-M and EGARCH will be presented. Their properties will be discussed in section 4. Methods to estimate these models will be presented and tests for the presence of ARCH will be discussed. Alternative approaches to modelling volatility and issues of temporal aggregation will be discussed as well. Sections 5 and 6 will be devoted to the use of ARCH- and GARCH-type models to describe the time variation of the risk premia in the forward foreign exchange market and in returns on stocks and bonds respectively. Section 7 concludes this chapter.

*The authors thank Frank de Jong and Peter Schotman for helpful comments on an earlier version of this chapter.
2. Volatility in Economic Models

In this section two models in which volatility plays a crucial role will be discussed. Section 2.1 deals with risk premia in forward foreign exchange. We discuss a simple example of the models for risk premia which have been developed building on Lucas (1982). Section 2.2 treats a capital asset pricing model with time-varying second moments.

2.1. Risk Premia in Foreign Exchange

Consider a model for two countries, $A$ and $B$, and two non-storable commodities, $X$ and $Y$. The endowments of the consumers of the countries $A$ and $B$ are $2X_t$ units of $X$ and zero units of $Y$, and $2Y_t$ units of $Y$ and zero units of $X$ respectively. Each consumer maximizes the expected time-additive utility over an infinite horizon:

$$E_0 \sum_{t=0}^{\infty} \beta^t U(X_t, Y_t)$$

where $\beta$ ($0 < \beta < 1$) is the time discount factor, and the subscript 0 indicates that the expectation is conditional on the information available in time period $t$ ($t = 0$). We assume a cash in advance economy. At the beginning of each period, per capita money balances in countries $A$ and $B$ are $M_t$ and $N_t$ respectively. Agents are assumed to engage in trade and to invest in assets after the uncertainty about the present state of the economy has been resolved.

We are interested in solving the model for the exchange rate $S_t$, the forward rate $F_t$, the forward risk premium and the interest rates $R^A_t$ and $R^B_t$ given the processes for the exogenous variables $X_t$, $Y_t$, $M_t$ and $N_t$.

Given the above endowments, prices of goods expressed in domestic currencies are $p_t^{XA} = 0.5 \frac{M_t}{X_t}$ and $p_t^{YB} = 0.5 \frac{N_t}{Y_t}$. The first-order conditions for expected utility maximization imply, for instance, for country $A$ that:

$$p_t^{YA}/p_t^{XA} = (\partial U_t/\partial Y_t)/(\partial U_t/\partial X_t)$$

where the r.h.s. is the marginal rate of substitution with $\partial U_t/\partial X_t = \partial U(X_t, Y_t)/\partial X_t$. The exchange rate $S_t$ expressing the value of the currency in $B$ in terms of that of $A$ can be written as $p_t^{YA} = p_t^{YB} \cdot S_t$. Substitution of this expression in (2.1.2) yields:

$$S_t = (p_t^{XA} \partial U_t/\partial Y_t)/(p_t^{YB} \partial U_t/\partial X_t)$$

Each unit of domestic currency which is saved has a yield after one period of $R^A_t$ and $R^B_t$ respectively.
The equilibrium conditions for postponed consumption are given by the following first-order condition for expected utility maximization:

$$R_t^A = \frac{\partial U_t}{\partial X_t} (p_t^{X_t})^{-1} / E_t [\beta \frac{\partial U_{t+1}}{\partial X_{t+1}} (p_{t+1}^{X_{t+1}})^{-1}]$$  \hspace{1cm} (2.1.4)

and an analogous expression for $R_t^B$. Finally, if there are no restrictions on the capital market, the covered interest rate parity leads to:

$$F_t = S_t R_t^A / R_t^B$$  \hspace{1cm} (2.1.5)

Along with Domowitz and Hakkio (1985), we assume that the utility function is of the Cobb–Douglas form:

$$U(X_t, Y_t) = X_t^\alpha Y_t^{1-\alpha} \quad (0 < \alpha < 1)$$  \hspace{1cm} (2.1.6)

and that the logarithm of the endowments with goods and money, denoted by the corresponding lower-case letters, are potentially conditionally heteroskedastic and generated by independent AR(1) processes. The autocorrelation coefficients are denoted by $\rho_x$, $\rho_y$, $\rho_m$ and $\rho_n$ respectively. The innovations in the log endowments, denoted by $u_{xt}$, $u_{yt}$, $u_{mt}$ and $u_{nt}$ are serially independent, normally distributed with mean vector zero and diagonal covariance matrix with diagonal elements $h_{xt}$, $h_{yt}$, $h_{mt}$ and $h_{nt}$. Substituting the expressions for the prices of goods and the marginal rate of substitution corresponding to (2.1.6) into (2.1.2) yields:

$$S_t = cM_t / N_t$$  \hspace{1cm} (2.1.7)

with $c = (1 - \alpha)/\alpha$. In order to find an expression for $R_t^A$, we write:

$$E_t[\frac{\partial U_{t+1}}{\partial X_{t+1}} (p_{t+1}^{X_{t+1}})^{-1}] = 2\alpha E_t[X_{t+1}^{\alpha} M_{t+1}^{-1} Y_{t+1}^{1-\alpha}]$$

$$= 2\alpha \exp[\alpha \rho_x x_t - \rho_m m_t + (1 - \alpha)\rho_y y_t$$

$$+ 0.5(\alpha^2 h_{xt} + h_{mt} + (1 - \alpha)^2 h_{yt})]$$  \hspace{1cm} (2.1.8)

Notice that we use the property that $E[\exp(x)] = \exp(\mu + 0.5\sigma^2)$ if $x \sim N(\mu, \sigma^2)$. Substituting (2.1.8) into (2.1.4) leads to:

$$R_t^A = \beta^{-1} \exp[\alpha(1 - \rho_x) x_t - (1 - \rho_m) m_t + (1 - \alpha)(1 - \rho_y) y_t$$

$$- 0.5\alpha^2 h_{xt} - 0.5h_{mt} - 0.5(1 - \alpha)^2 h_{yt}]$$  \hspace{1cm} (2.1.9)

For $R_t^B$, the same expression as (2.1.9) can be obtained with the exception that $\rho_m$, $m_t$ and $h_{mt}$ have to be replaced by $\rho_n$, $n_t$ and $h_{nt}$ respectively. Substituting
the expression for $R_t^A$ and $R_t^B$ into (2.1.5), the logarithm of the forward rate, $f_t$, can be expressed as:

$$f_t = \ln \left( c + \rho_m m_t - \rho_n n_t - 0.5h_{mt} + 0.5h_{nt} \right)$$  \hspace{1cm} (2.1.10)

If similarly $s_t$ denotes the log spot rate, (2.1.7) yields:

$$E_t s_{t+1} = \ln \left( c + \rho_m m_t - \rho_n n_t \right)$$  \hspace{1cm} (2.1.11)

so that the forward risk premium defined as $E_t s_{t+1} - f_t$ becomes:

$$E_t s_{t+1} - f_t = 0.5(h_{mt} - h_{nt})$$  \hspace{1cm} (2.1.12)

The forward risk premium depends on the difference between the conditional variances of the money balances in the two countries. If the uncertainty about money balances is the same in both countries, the risk premium is zero, i.e. the forward rate is an unbiased forecast of the future spot rate, and uncovered interest rate parity will hold as well.

Most importantly, the simple model shows that conditional variances of economic variables play a crucial role in economic models that take uncertainty into account. If these variances vary through time, it will be crucial to appropriately model the time dependence of $h_{mt}$ and $h_{nt}$ in order to explain the behaviour of the forward premium. Along the same lines the more sophisticated model of Hodrick (1989) leads to an expression in which the variances of income and the shares of government expenditures become additional determinants of the risk premium.

### 2.2. The Static Capital Asset Pricing Model

In this subsection we consider a second example of a model in which volatility plays a crucial role. We postulate a representative investor, maximizing a utility function defined over the expected value and the variance of end-of-period wealth $W_{t+1}$:

$$\max \ U[E_t(W_{t+1}), \sigma_t^2(W_{t+1})]$$  \hspace{1cm} (2.2.1)

where

$$E_t(W_{t+1}) = W_t x_t^\top E_t(R_{t+1}) + W_t(1 - x_t^\top)R_t^f$$  \hspace{1cm} (2.2.2)

$$\sigma_t^2(W_{t+1}) = W_t^2 x_t^\top \Sigma_t x_t$$  \hspace{1cm} (2.2.3)

where $W_t$ represents the investor's wealth and $x_t$ is an $n \times 1$ vector of investment shares in risky assets whose rates of return have conditional means and
covariances denoted by $E_t(R_{t+1})$ and $\Sigma_t$ respectively. $R_t^f$ is the rate of return on a risk-free asset, and $\iota$ is a unit vector.

The first-order conditions for the maximization problem (2.2.1) yield:

$$x_t = (\rho_t \Sigma_t)^{-1} (E_t(R_{t+1}) - \iota R_t^f)$$  

(2.2.4)

where $\rho_t$ is the relative risk aversion coefficient, $\rho_t = -2W_t^1U_2/U_1$, and $U_1$ and $U_2$ are the partial derivatives of $U$ with respect to the first and second argument in (2.2.1) respectively (assumed to be $U_1 > 0$, $U_2 < 0$). Equation (2.2.4) determines the optimal composition of the portfolio of the investor, the determinants of which are $\rho_t$, $R_t^f$ and the expected returns and the covariance matrix of the risky assets. Equation (2.2.4) can also be solved for the equilibrium expected returns:

$$E_t(R_{t+1}) - \iota R_t^f = \rho_t \Sigma_t x_t = (-2U_2/U_1) \Sigma_t W_t x_t$$  

(2.2.5)

where $W_t x_t$ is the actual value which in equilibrium is equal to demand and supply. Since the expectation of $R_{t+1}$ equals its realization minus the forecast error $\xi_{t+1}$, we have:

$$R_{t+1} = \iota R_t^f + \rho_t \Sigma_t x_t + \xi_{t+1}$$  

(2.2.6)

where $\xi_{t+1}$ has conditional mean zero and conditional covariance matrix $\Sigma_t$.

Model (2.2.6) has the property that there are restrictions between the conditional mean of future returns and their conditional covariance matrix. Notice that (2.2.5) can also be expressed in terms of the familiar $\beta$ coefficients. From (2.2.5), the expected excess return of the portfolio with shares $x_t$ is given by $\rho_t x_t' \Sigma_t x_t$. Therefore when we substitute this expression into (2.2.5), we get the capital asset pricing model (CAPM),

$$E_t R_{t+1} - \iota R_t^f = \beta_t (E_t R_{t+1}^p - R_t^f)$$  

(2.2.7)

where $\beta_t = x_t' \Sigma_t x_t / \text{var}_t(R_{t+1}^p)$ and $R_t^p$ denotes the return on the portfolio. Note that $\beta_t$ is not assumed to be time-invariant.

As in the previous subsection, the covariance matrix $\Sigma_t$ plays a central role in the model. If $\Sigma_t$ is time-dependent, it will be crucial for the analysis of the CAPM (2.2.6) to get a good specification for this time variation.

2.3. Concluding Remarks

The two examples have a basic feature in common. Volatility of the series measured by the conditional variances plays an important role in the economic explanation of the level (or the conditional mean) of variables such as the
forward risk premium in exchange markets and expected future returns in asset markets. In other words, the conditional mean is linearly related to the conditional variance. Similar models have also been used in modelling the pricing of futures. For a recent application, we refer to Nijman and Beetsma (1991) and Baillie and Myers (1991). Any time-dependence of the conditional variance therefore has direct implications for the time-dependence of the conditional mean. In the next section, we shall present models which are designed to describe the time-variation of conditional variances.

3. Models for Conditional Heteroskedasticity

Econometric models with time-varying conditional variances have recently received much attention in the literature. In a seminal paper, Engle (1982) introduced stochastic models of the form:

\[ y_t = \epsilon_t h_t^{1/2} \quad (3.1) \]

\[ h_t = \alpha_0 + \alpha_1 y_{t-1}^2 + \cdots + \alpha_q y_{t-q}^2, \quad \alpha_0 > 0, \alpha_i \geq 0, \sum_{i=1}^{q} \alpha_i < 1 \quad (3.2) \]

where the \( \epsilon_t \) are i.i.d. with \( E(\epsilon_t) = 0 \) and \( \text{var}(\epsilon_t) = 1 \). This is a \( q \)th order autoregressive conditional heteroskedasticity (ARCH) model. Adding the assumption of conditional normality, the model can be written as:

\[ y_t \mid \Phi_{t-1} \sim N(0, h_t) \quad (3.3) \]

with \( h_t \) being given by (3.2) and \( \Phi_{t-1} \) being the set of information available at time \( t - 1 \). The non-negativity of the \( \alpha_i \)s is required for the variance to be non-negative, whereas the requirement that the \( \alpha_i \)s sum to less than one is needed for \( y_t \) to be wide sense stationary (see Engle (1982)). In the first order ARCH-model, the conditional variance of \( y_t \) will increase when \( y_{t-1}^2 \) increases and decrease as \( y_{t-1}^2 \) decreases. The ARCH-model is a generalization of linear models with homoskedastic disturbances in which the conditional mean varies with the variables in \( \Phi_{t-1} \) (such as ARMA-models or linear regression models), for which the conditional variance is constant across time. It is closely related to the bilinear models introduced by Granger and Andersen (1978) (see also Weiss (1986a)), an example of which is of the form \( y_t = \epsilon_t y_{t-1} \), with conditional variance \( \sigma^2 y_{t-1}^2 \) so that the unconditional variance becomes zero, one or infinity, depending on the value of \( \sigma^2 \).
A generalization proposed by Bollerslev (1986) is the GARCH model. For the GARCH(1, 1) model, the conditional variance is given by:

\[ h_t = \alpha_0 + \alpha_1 y^2_{t-1} + \beta h_{t-1} \]  

(3.4)

with \( \alpha_0, \alpha_1, \beta > 0, \alpha_1 + \beta < 1 \). Equation (3.4) can be written as

\[ h_t = \sigma^2 + \alpha_1 \sum_{i=0}^{\infty} \beta^i (y^2_{t-i-1} - \sigma^2) \]  

(3.5)

with \( \sigma^2 = \alpha_0(1 - \alpha_1 - \beta) \). According to the assumption made in (3.4) (or (3.5)), the conditional variance of \( y_t \) will be large if the weighted average of past \( y^2_t \) with geometrically declining weights is large. When \( y_t \) is stationary, \( E y^2_t = E y^2_{t-k}, (3.5) \) implies that \( E y^2_t = E(E(y^2_t | \Phi_{t-1})) = Eh^2_t = \sigma^2 \) so that the conditional variance (3.5) can become larger than the unconditional variance if past realizations of \( y^2_t \) have been larger than \( \sigma^2 \). The realizations of a GARCH(1, 1) model exhibit clusters of large values. A comparison of the GARCH(1, 1) model (3.4) with the ARCH(q) model (3.3) indicates that the former one may be seen as a parsimonious parametrization with features similar to those of the ARCH(q) models with exponentially declining coefficients \( \alpha_i \). Notice the similarity between the ARMA(1, 1) model and a high order MA(q) process.

Of course, along with Bollerslev (1986), one can also consider the extension of (3.4) to the GARCH(p, q) model for the conditional variance of (3.1):

\[ h_t = \alpha_0 + \sum_{i=1}^{p} \beta_i h_{t-i} + \sum_{i=1}^{q} \alpha_i y^2_{t-i} \]  

(3.6)

with \( \alpha_0, \beta_i, \alpha_i > 0, \Sigma_i \beta_i + \Sigma_i \alpha_i < 1 \). Again, the non-negativity conditions imply a non-negative variance, while the condition that the sum of the \( \alpha_i \)s and \( \beta_i \)s is smaller than one is required for wide sense stationarity of \( y_t \) (see Bollerslev (1986)).

When \( \Sigma_i \beta_i + \Sigma_i \alpha_i = 1 \), the integrated GARCH (IGARCH) model arises (see Engle and Bollerslev (1986)). It has a unit root in the autoregressive polynomial of the variance function. For instance, for the simple IGARCH(1, 1) model with \( \alpha + \beta = 1 \), the minimum mean square error forecast for the conditional variance \( s \) steps ahead, \( h_{t+s} \), is equal to \( \alpha_0(s - 1) + h_{t+1} \).

As for the random walk model for the series itself, current information remains important for forecasts of the conditional variance for all horizons in the IGARCH(1, 1) model. Although there are similarities with the model that is integrated in the mean, the problems which arise if one estimates a model with unit roots in the conditional mean do not arise in this case (see e.g. Lumsdaine (1990)).

In the empirical analysis of economic data, higher-order GARCH models do not generally yield a better representation of the features of the data than the
GARCH(1, 1) or the GARCH(1, 2) models. A similar finding holds for the choice of the order of ARMA models to describe the serial correlation of many univariate macroeconomic series.

For all the models introduced above, the reaction of the conditional variance is symmetric for increases as well as decreases of the same size of the variables in $\Phi_{t-1}$. A specification which allows for asymmetric reactions of the conditional variance is the exponential GARCH (EGARCH) model put forward by Nelson (1991):

$$\ln h_t = \alpha_0 + \sum_{i=1}^{p} \beta_i \ln h_{t-i} + \sum_{i=1}^{q} \alpha_i (\varphi \varepsilon_{t-i} + \gamma |\varepsilon_{t-i}| - \gamma E|\varepsilon_{t-i}|)$$  \hspace{1cm} (3.7)

where the parameters $\alpha_0$, $\beta_i$ and $\alpha_i$ are not restricted to being non-negative. An asymmetric reaction in the variance of the return of stocks can be expected as a result of the so-called leverage effect. A negative shock to the returns increases the debt-equity ratio and therefore increases uncertainty of future returns. In the EGARCH(1, 1) model, this could occur when $\alpha_1 > 0$ and $\varphi < 0$.

Engle, Lilien and Robins (1987) introduced the ARCH-M model in which the conditional mean is a function of the conditional variance of the process

$$y_t = g(x_{t-1}, h_t) + \varepsilon_t h_t^{1/2}$$ \hspace{1cm} (3.8)

where $x_{t-1}$ is a vector of predetermined variables, $g$ is some function of $x_{t-1}$ and $h_t$, and $h_t$ is generated by an ARCH($q$) process. The most common ARCH-M model simply has $g(x_{t-1}, h_t) = \delta h_t$. As the examples presented in section 2 show, many theories in finance involve an explicit trade-off between risk and expected return, leading to models in which the conditional mean depends on the conditional variance.

The models presented until now have all been univariate. As indicated by the theoretical model in section 2.2, the analysis of many issues of asset pricing and portfolio allocation requires a multivariate framework. Extensions of the models presented in this section to a multivariate setting will be considered in sections 5 and 6.

Measures of volatility which are not based on ARCH-type models have also been put forward in the literature. For instance, French et al. (1987) construct monthly stock return variance estimates by taking the average of the squared daily returns. To assess the temporal dependence, standard time series models are subsequently estimated for these variance estimates. This procedure does not make efficient use of all the data. Another drawback of this approach is that it does not yield the high frequency variance forecasts which are required in many models originating in financial theory. Furthermore, the conventional standard errors from the second-stage estimation may not be appropriate. Nevertheless, the computational simplicity of this procedure and a related model put forward by Schwert (1989), in which the conditional standard deviation is measured by the absolute value of the residuals from a first step estimate.
of the conditional mean, makes them appealing alternatives to more complicated ARCH type models for preliminary data analysis.


In this section, we shall summarize the main results about the statistical properties of the models presented in the preceding section and we shall present estimation and testing procedures for these models. Finally, we shall discuss some issues related to temporal aggregation of ARCH and GARCH models.

4.1. Moments and Stationarity

Bollerslev (1986) has shown that under normality the GARCH process defined in (3.6) is wide sense stationary with \( E(y_t) = 0 \) and \( \text{var}(y_t) = \alpha_0(1 - \alpha(1) - \beta(1))^{-1} \) and \( \text{cov}(y_t, y_s) = 0 \) for \( t \neq s \), if and only if \( \alpha(1) + \beta(1) < 1 \). For the GARCH(1,1) model given by (3.6) when \( p = q = 1 \), a necessary and sufficient condition for the existence of the 2\( m \)th moment is \( \Sigma_{j=0}^{m} (j!) a_j \alpha^j \beta^{m-j} < 1 \), where \( a_0 = 1 \) and \( a_j = \prod_{i=1}^{j} (2i - 1) \), \( j = 1, 2, \ldots \). Finally, Bollerslev (1986) has given a recursive formula for the even moments of \( y_t \) when \( p = q = 1 \). The fourth moment of a conditionally normal GARCH(1,1) variable e.g. will be \( E y_t^4 = 3(E y_t^2)^2 [1 - (\beta + \alpha_1)^2] / [1 - (\beta + \alpha_1)^2 - 2\alpha_1^2] \) if this moment exists. As a result of the symmetry of the normal distribution, the odd moments are all zero if they exist.

4.2. Estimation of GARCH Models

The parameters of the models described here can be estimated by the maximum likelihood (ML) method. For simplicity, we discuss estimation of the parameters of the GARCH(1,1) model under the assumption that \( \epsilon_t \sim \mathcal{N}(0,1) \). The log-likelihood function for a sample of \( T \) observations \( y_1, y_2, \ldots, y_T \), can be written as

\[
L(y_1, y_2, \ldots, y_T \mid \vartheta) = \sum_{t=1}^{T} \left[ c - \frac{1}{2} \ln h_t(\vartheta) - \frac{1}{2} y_t^2 / h_t(\vartheta) \right] \tag{4.2.1}
\]

where \( \vartheta = (\alpha_0, \alpha_1, \beta) \), \( h_1(\vartheta) = \sigma^2 = \alpha_0(1 - \alpha_1 - \beta) \) and \( h_t(\vartheta) = \alpha_0 + \beta h_{t-1}(\vartheta) + \alpha_1 y_{t-1}^2 \) (\( t > 1 \)). Given (initial) values of the parameters \( \vartheta \), the value of the log-likelihood function (4.2.1) can be recursively evaluated. Compute \( h_1(\vartheta) \), then get \( h_2(\vartheta) \) etc. and substitute the values into (4.2.1) to get the value of \( L \). Standard numerical procedures can be used to compute the maximum of \( L \) in (4.2.1).
Under regularity conditions which are given in e.g. Crowder (1976), the value of which \( \theta \) maximizes \( L \), \( \hat{\delta}_{ML} \), is consistent, asymptotically normally distributed and efficient,

\[
\sqrt{T}(\hat{\delta}_{ML} - \theta) \overset{d}{\sim} N(0, \text{var}(\hat{\delta}_{ML}))
\]

(4.2.2)

with \( \text{var}(\hat{\delta}_{ML}) = [E(T^{-1} \partial^2 L/\partial \theta \partial \theta')]^{-1} \). The asymptotic covariance matrix of \( \hat{\delta}_{ML} \) can be consistently estimated by computing the inverse of the Hessian matrix associated with the log-likelihood function (4.2.1) evaluated at \( \hat{\delta}_{ML} \). Most authors assume that the fourth moment of the data exists. Conditions for the existence of \( E\gamma_t^4 \) have been discussed in section 4.1. Under that assumption the proof of consistency of the estimators is greatly simplified. We note that Lumsdaine (1990) proves consistency without requiring fourth moments to exist.

The estimation of the ARCH-M model poses no extra problems. However, as shown by Engle (1982), in absence of ARCH-M effects, the information matrix for the model under conditional normality is block-diagonal between the parameters in the conditional mean and the variance functions. This is no longer true for the ARCH-M model. Thus unlike the ARCH model where consistent estimates of the parameters in the conditional mean, \( g(x_{t-1}) \), can be obtained even if \( h_t \) is misspecified, consistent estimation in the ARCH-M model requires the full model to be correctly specified.

Finally, if the parameters of the conditional variance in a regression model are nuisance parameters only, one can of course simply estimate the parameters in the conditional mean by ordinary least squares if the conditional variance does not affect the conditional mean. Note, however, that one has to use heteroskedasticity consistent standard errors, as proposed by White (1980), to conduct valid inference. Suppose e.g. that \( \rho \) is the parameter of interest in the regression \( y_t = \rho y_{t-1} + u_t \) where \( u_t \) is uncorrelated with mean zero and conditional variance \( \alpha_0 + \alpha_1 u_{t-1}^2 \), i.e. the error is ARCH(1). The ordinary least squares estimator of \( \rho \) is a pseudo maximum likelihood estimator and its large sample distribution is given by \( \sqrt{T}(\hat{\rho}_{OLS} - \rho) \overset{d}{\sim} N(0, A^{-1}BA^{-1}) \) where \( A = E[-T^{-1} \partial^2 C/\partial \rho^2] \) and \( B = E[T^{-1} (\partial C/\partial \rho)^2] \) and \( C \) is the pseudo log likelihood \( C = -\sum_{t=1}^{T} (y_t - \rho y_{t-1})^2 \), see e.g. Gouriéroux et al. (1984). The matrices \( A \) and \( B \) can easily be estimated consistently.

One of the implications of the need to correct the standard errors of estimators of parameters in the conditional mean is that in the presence of ARCH, standard tests for serial correlation in \( y_t \) will be upward biased, thus leading to over-rejections. Consider e.g. a test for first-order autocorrelation if the data are in fact generated by an ARCH(1) model and assume conditional normality. The standard estimate of the first-order autocorrelation coefficient coincides with the ordinary least squares estimator in a first-order auto-regressive model. From the results above it is easily checked that under the null hypothesis that \( \rho = 0 \) the large sample variance of \( \hat{\rho}_{OLS} \) will be \( T^{-1}\{\alpha_0 E\gamma_{t-1}^2 + \alpha_1 E\gamma_{t-1}^2\}/\{E\gamma_{t-1}^2\}^2 \) if fourth moments exist. From section 4.1
this can be rewritten as \( T^{-1}(1 + 2\alpha_1 - 3\alpha_1^2)(1 - 3\alpha_1^2)^{-1} \), when \( \alpha_1 < 1/\sqrt{3} \). For \( \alpha_1 = 0.5 \), this large sample variance exceeds the conventional asymptotic variance, \( 1/T \), by a factor of five (see e.g. Diebold (1987) for further discussion).

## 4.3. Testing for Conditional Heteroskedasticity

The likelihood ratio (LR) criterion can be used to test the hypothesis of conditional homoskedasticity e.g. against the GARCH(1, 1) alternative in (3.4). The LR-statistic associated with \( H_0: \alpha_1 = 0 \) and \( \beta = 0 \) does not have a \( \chi^2 \)-distribution with two degrees of freedom in this case, as the alternative hypothesis is \( H_1: \alpha_1 > 0 \) and \( \beta > 0 \). The standard assumption that the true parameter value under \( H_0 \) does not lie on the border of the parameter space under \( H_1 \) does not hold. An LR test which uses a \( \chi^2 \)-distribution with two degrees of freedom can be shown to be conservative (see e.g. Kodde and Palm (1986)), that is if it rejects the null hypothesis, then the LR-test which uses the correct distribution will certainly reject the null hypothesis. Demos and Sentana (1990) present critical values for the LR and Wald tests for testing ARCH effects versus constancy of the variance of a series. A second problem arises as the parameter \( \beta \) is not identified if \( \alpha_1 = 0 \). Therefore a test of \( H_0: \alpha_1 = 0 \) could yield misleading results. Engle (1984) shows how to carry out a test if some nuisance parameters are not identified under the null hypothesis.

A simple and frequently used test for conditional heteroskedasticity is the LM test of the hypothesis \( H_0: \alpha_1 = \alpha_2 = \ldots = \alpha_q = 0 \) in the ARCH(q) model in (3.2). It has the form:

\[
\text{LM} = (\partial L/\partial \alpha') (\partial^2 L/\partial \alpha \partial \alpha')^{-1} (\partial L/\partial \alpha) |_{\alpha = \hat{\alpha}}
\]

(4.3.1)

where \( \hat{\alpha} \) denotes the ML estimate of \( \alpha \) under \( H_0 \). For the ARCH(q) model, \( h_t = z_t' \alpha \) and

\[
(\partial L/\partial \alpha)_{\alpha = \hat{\alpha}} = (1/2\hat{\alpha}_0) \sum_t z_t (y_t^2/\hat{\alpha}_0 - 1) = (1/2\hat{\alpha}_0) z' f_0
\]

(4.3.2)

where \( z_t' = (1, y_{t-1}^2, \ldots, y_{t-q}^2) \) and \( z = (z_1, z_2, \ldots, z_T) \) and \( f_0 \) is the column vector of \( (y_T^2/\hat{\alpha}_0 - 1) \).

The Hessian matrix can be written as:

\[
J_{\alpha \alpha} |_{\alpha = \hat{\alpha}} = (1/2\hat{\alpha}_0) z' z
\]

(4.3.3)

and therefore the LM test statistic can be written as:

\[
\text{LM} = 1/2 f_0' z (z' z)^{-1} z' f_0
\]

(4.3.4)
Notice that \( \text{plim} \frac{f_0 f_0}{T} = 2 \), because normality has been assumed. Therefore, an asymptotically equivalent statistic would be:

\[
\text{LM} = T f_0 z (x' z)^{-1} z f_0 f_0 = TR^2
\]

(4.3.5)

where \( R^2 \) is the squared multiple correlation between \( f_0 \) and \( z \). Since adding a constant and multiplying by a scalar will not change the \( R^2 \) of a regression, this is also the \( R^2 \) of a regression of \( y_t^2 \) on an intercept and \( q \) lagged values of \( y_t^2 \). Demos and Sentana (1990) propose a simple one-sided version of the \( TR^2 \)-type LM test for ARCH in (4.3.5), which is computed from the same auxiliary regression of the squares of the residuals on a constant and its lags. They report critical values for the one-sided LM test. These critical values are robust to non-normality.

### 4.4. Non-Normality of the Conditional Density of \( \varepsilon_t \)

As discussed above, financial return series often exhibit volatility clustering, and their unconditional distribution tends to have fatter tails than the normal distribution. Moreover, the conditional distribution of the standardized variable \( \frac{y_t}{\sqrt{h_t}} \) often appears to be leptokurtic.

Bollerslev (1987) suggests use of the standardized Student-\( t \) distribution with the degrees of freedom \( \nu \) being estimated. The log-likelihood of this model is:

\[
L = T \ln \ k(\nu) - \sum_{t=1}^{T} 0.5(\nu + 1)\ln [1 + (\nu - 2)^{-1} e_t^2 / h_t] \quad (4.4.1)
\]

with \( k(\nu) = \Gamma((\nu + 1)/2) \Gamma(\nu/2)^{-1}(\pi(\nu - 2))^{-1/2} \). The \( t \)-distribution has fatter tails than the normal distribution. As \( \nu \) increases, it converges to the normal distribution. Other densities that have been used in the estimation of ARCH models are the normal-Poisson mixture distribution (see e.g. Jorion (1988), Nieuwland et al. (1991)) and the normal-lognormal mixture distribution (see e.g. Hsieh (1989)). An alternative to parametric ML estimation is the use of a semi-parametric density estimation technique which approximates the density function (see e.g. Gallant and Tauchen (1989) or Engle and Gonzalez-Rivera (1991)). Notice that semi-parametric and non-parametric methods can also be used to approximate an unknown conditional variance function (see Gallant and Tauchen (1989) and Pagan and Ullah (1988)).

In applications, the distribution of the disturbance is often unknown. Weiss (1986b) and Bollerslev and Wooldridge (1990) have shown that quasi-maximum likelihood estimators which are based on a normality assumption are consistent under weak assumptions even if the normality assumption does not hold. Moreover they show how asymptotic standard errors for the parameters in the conditional mean and variance functions that are robust to departures from normality can be obtained.
4.5. Temporal Aggregation of ARCH Processes

Issues of temporal aggregation play an important role in time series modelling, in particular when the investigator has the choice between using data observed with a high frequency (e.g. daily observations) or using observations (e.g. monthly) sampled less frequently.

In this section, we illustrate how temporal aggregation affects the structure of the model in the case where the high frequency data are generated by an ARCH(1) model. This section draws on Drost and Nijman (1990) who also consider more general cases. Diebold (1988) has shown that the conditional heteroskedasticity disappears in the limit if the sampling frequency decreases and that in case of low variables the implied marginal low frequency distribution converges to the normal distribution. Nelson (1990) considered the case of an increasing sampling frequency and derived the limiting continuous time model.

We consider an ARCH(1) model for a stock variable $y_t$ which is observed every second period ($t = 2, 4, \ldots T$), with, for simplicity reason, $T$ being assumed to be even. Along with Drost and Nijman (1990) we distinguish between three forms of ARCH models. We say that $y_t = \varepsilon_t h_t^{1/2}$ is generated by a semi-strong ARCH(1) model if:

$$E(y_t | y_{t-1}, y_{t-2}, \ldots) = 0 \quad (4.5.1)$$

$$E(y_t^2 | y_{t-1}, y_{t-2}, \ldots) = h_t \quad (4.5.2)$$

and $h_t$ is generated by (3.2) with $q = 1$. The strong form corresponds to the case where the disturbance is assumed to be identically distributed with mean zero and variance equal to one, while for the weak form, condition (4.5.1) holds but condition (4.5.2) is replaced by the condition that the projection of $y_t^2$ on the space generated by $y_{t-1}, y_{t-2}, \ldots$ and its squares is equal to $h_t$.

Integrating expressions (4.5.1) and (4.5.2) with respect to $y_{t-1}, y_{t-3}, \ldots$ we get for the low frequency model:

$$E(y_t | y_{t-2}, y_{t-4}, \ldots) = \bar{\alpha}_0 + \bar{\alpha}_1 y_{t-2}^2 = \bar{h}_t \quad (4.5.4)$$

with $\bar{\alpha}_0 = \alpha_0 (1 + \alpha_1)$ and $\bar{\alpha}_1 = \alpha_1^2$. The low-frequency model is semi-strong ARCH(1) as well, with parameters $\bar{\alpha}_0$ and $\bar{\alpha}_1$ replacing the high frequency parameters $\alpha_0$ and $\alpha_1$.

A natural question to consider is whether the class of strong ARCH(1) models for stock variables is closed under temporal aggregation as well.
Assuming that $\varepsilon_t$ is i.i.d. with distribution $D(0, 1)$, the rescaled disturbances in the low-frequency model are defined by

$$\nu_t = y_t / \tilde{h}_t^{1/2}$$  \hspace{1cm} (4.5.5)

Rewriting (4.5.5), one can easily show that:

$$\nu_t = \varepsilon_t [\lambda_t + (1 - \lambda_t) \varepsilon_{t-1}^2]^{1/2} \text{ with } \lambda_t = \tilde{\alpha}_0 \tilde{h}_t^{1/2}$$  \hspace{1cm} (4.5.6)

According to (4.5.6) the rescaled disturbances in the low-frequency model depend on past observations, even if the rescaled disturbances in the high frequency model are i.i.d. From (4.5.6), one can show that:

$$E(\nu_t^4 | y_{t-2}, y_{t-4}, ...) = \kappa (\kappa - 1) (\lambda_t - 1)^2 + \kappa$$  \hspace{1cm} (4.5.7)

where $\kappa = E\varepsilon_t^4$ which implies that the fourth moment of the low-frequency conditional distribution of the rescaled disturbances will depend on the information set in a way similar to the dependence of the second moments on the past in the GARCH model. The low-frequency model that is implied by a high-frequency strong ARCH(1) process is no longer strong ARCH(1), although it still satisfies the assumptions of semi-strong ARCH(1) as shown above. Therefore, the assumption that rescaled disturbances are i.i.d. at the frequency at which the data happen to be available is arbitrary. More detailed results for higher-order ARCH and for GARCH models for stock and flow variables can be found in Drost and Nijman (1990). In general, it can be said that aggregating strong GARCH processes does not lead to a strong GARCH model. Weak GARCH(1, 1), defined above, generally leads to weak GARCH in the temporally aggregated data.

4.6. Concluding Remarks

In this section we have discussed statistical procedures that can be used to handle GARCH models. Conditional heteroskedasticity of high-frequency financial series is by now a well-established stylized fact. Fama (1965) already observed that price changes tended to be dependent over time and characterized by tranquil and volatile periods, and that the unconditional distributions of the price changes were typically fat-tailed and leptokurtic. Recently, Baillie and Bollerslev (1989) and Hsieh (1989) carried out an extensive study of the time series properties of exchange rates. They show that the first differences of the logarithms of the daily rates are approximately uncorrelated through time, and a GARCH(1, 1) model with near unit roots and conditionally t-distributed errors is found to be a good representation to the leptokurtosis and time-dependent conditional heteroskedasticity. The parameter estimates are
similar for the different currencies. The results carry over to weekly and monthly data in which, in line with the results in section 4.5, the degree of time-dependent heteroskedasticity is reduced if the length of the sampling interval increases. Related stylized facts for stock returns and interest rates can be found in Bollerslev et al. (1992).

5. Econometric Models of Time-Varying Risk Premia in Foreign Exchange

5.1. Introduction

As shown in section 2.1, Domowitz and Hakkio (1985) (DH85 from now on) have developed a simple model for the risk premium component in forward foreign exchange rates, using the general set-up in Lucas (1982). In this section various ways of testing their model will be described. In section 5.2 we will describe the empirical results in DH85 which are based on a univariate ARCH-M model and monthly data. Section 5.3 will be devoted to Baillie and Bollerslev (1990) where results from a multivariate GARCH-M and weekly overlapping data are provided. Finally section 5.4 is devoted to non-parametric estimation of the variance functions as proposed by Pagan and Ullah (1988). Before we start the discussion of direct tests of the model, it is worthwhile to refer to some related work which is available in the literature.

The existence of a risk premium in forward foreign exchange has been claimed by many authors (see e.g. Baillie, Lippens and McMahon (1983) and Hansen and Hodrick (1980)) who tested the unbiasedness of the forward rate. In a second line of research several authors have tried to assess the importance of time variation in risk premia, without making the strong structural assumptions in DH85. Fama (1984) and Hodrick and Srivastava (1986) have shown that the unconditional variance of the risk premium is greater than the unconditional variance of the expected rate of depreciation. Similarly Wolff (1987) and Nijman, Palm and Wolff (1991) have presented evidence on the conditional variances of the risk premia and expected rate of depreciation using weak structural assumptions only. Wolff (1987) has moreover shown how estimates of the premium can be obtained under similar assumptions. Finally, in a third line of research, Hansen and Hodrick (1983) and Hodrick and Srivastava (1984) have tested the validity of the Euler equation (2.1.4). Evidently the results from these lines of research are taken as starting points in the literature on structural models for the risk premium to be discussed in sections 5.2 to 5.4. An excellent discussion of the literature on these issues which was available about five years ago is provided in Hodrick (1987).

In section 2.1 we outlined how DH85 derive the model

$$s_{t+1} - f_t = \delta \{ h_{nt+1} - h_{nt+1} \} + u_{nt+1} - u_{nt+1}$$  \hspace{1cm} (5.2.1)

$$u_{nt+1} \mid I_t \sim N(0, h_{nt+1}), \hspace{1cm} u_{nt+1} \mid I_t \sim N(0, h_{nt+1})$$

where $\delta = 0.5$. Equation (5.2.1) can be estimated jointly with the two money supply equations, and along these lines the hypothesis $\delta = 0.5$ can be tested. This approach will be discussed in section 5.4. Instead, DH85 approximated (5.2.1) by:

$$s_{t+1} - f_t = \mu + \delta h_{t+1} + e_{t+1}$$ \hspace{1cm} (5.2.2)

$$e_{t+1} \mid I_t \sim N(0, h_{t+1})$$

where $h_{t+1}$ is assumed to be ARCH(4).

Estimation of the model in (5.2.2) only requires spot and forward exchange rate data and is more straightforward than the multivariate approach which models the money supplies as well. On the other hand, (5.2.2) is only implied by (5.2.1) in special cases such as when one of the two money supplies is conditionally homoskedastic and an ARCH(4) model holds for the conditional variance of the money supply in the other country. Nevertheless, one might expect some impact of $h_t$ on $s_{t+1} - f_t$ in more general cases as well. Note that the estimated risk premium can change sign because of the presence of the constant term.

No doubt motivated by the literature on testing for unbiasedness of the forward rate, which suggested a relation between $s_{t+1} - s_t$ and $f_t - s_t$, DH85 finally include the forward premium $f_t - s_t$ as a regressor by adding $f_t - s_t$ to both sides of the equality which yields their final specification:

$$s_{t+1} - s_t = \mu + \gamma (f_t - s_t) + \delta h_{t+1} + e_{t+1}$$ \hspace{1cm} (5.2.3)

$$e_{t+1} \mid I_t \sim N(0, h_{t+1})$$

DH85 have estimated (5.2.3) by maximum likelihood, assuming that $h_{t+1}$ is generated by an ARCH(4) process, as has been discussed in (4.2.1) for a related case. The standard errors of the estimators have been obtained using a variant of (4.2.2). For monthly data on exchange rates of the pound sterling, the French franc, the German mark and the Swiss franc against the US dollar from June 1973 to September 1982 the results indicate that all four series show very little conditional heteroskedasticity, which with hindsight is not so surprising, given the results in Baillie and Bollerslev (1989) and Drost and
The hypothesis that \( \mu = 0, \gamma = 1 \) and \( \delta = 0 \) can be rejected for the pound sterling, but could not be rejected for the other three currencies. DH85 therefore conclude that ‘there is little support for the conditional variance of the exchange rate forecast error being an important sole determinant of the risk premium’.

5.3. The Weekly Multivariate GARCH-M Model for the Risk Premium in Baillie and Bollerslev (1990)

Baillie and Bollerslev (1990) (BB90 from now on) have extended the work of DH85 in at least three ways. First of all they analyse exchange rates of a number of currencies against the dollar in a multivariate setting, thereby avoiding the approximation of (5.2.1) by (5.2.2). Second, they replace the ARCH(4) assumption on the variances by a GARCH(1,1) assumption, which is more parsimonious and is less likely to yield a priori implausible parameter estimates. Finally BB90 analyse weekly data, which contain much more conditional heteroskedasticity than the monthly data used in DH85 and are therefore more likely to reveal a relation between time-varying variances and time-varying risk premia. As no weekly forward contracts are traded, however, they have to stick to monthly forward rates, which creates the problem of overlapping samples discussed before, e.g. in Hansen and Hodrick (1980). In this section the three extensions referred to above will be added to the model one after another, and some of the empirical results in BB90 will be discussed.

Starting with the extension to exchange rates of a number of currencies against the dollar, note that DH85 assume that the model which has been derived holds for all bilateral exchange rates. Note, however, that the underlying model describes a two-country world. Extension to a multi-country setting without affecting the results appears to require extreme separation assumptions, a topic that will not be pursued here. Consider an \( N+1 \) country world \((i = 0, \ldots, N)\) where the money stocks are denoted by \(m_{i,t}\) and where all exchange rates \(s_{i,t} (i = 1, \ldots, N)\) are denominated in the currency of the first country \((i = 0)\). Equation (2.1.7) of the DH85 model implies that

\[
\text{var}_t[s_{i,t+1}] = \text{var}_t[m_{0,t+1}] + \text{var}_t[m_{i,t+1}] \quad (5.3.1)
\]

and

\[
\text{cov}_t[s_{i,t+1}, s_{j,t+1}] = \text{var}_t[m_{0,t+1}], \quad i \neq j \quad (5.3.2)
\]

because of the assumed independence of the AR(1) processes which generate the money supplies. Substitution of (5.3.1) and (5.3.2) in (2.1.12) yields:

\[
E_t[s_{i,t+1}] - f_{i,t} = \text{cov}_t[s_{i,t+1}, s_{j,t+1}] - 0.5 \text{var}_t[s_{i,t+1}] \quad (5.3.3)
\]

which is an expression for the risk premium which avoids the use of money
supply data and which can be tested in the context of a multivariate GARCH-M model. The model which BB90 propose for an \( N+1 \) currency world is:

\[
    s_{i,t+1} - f_{i,t} = \mu_i + \sum_{j=1}^{N} \delta_{ij} \text{cov}_t(s_{i,t+1}, s_{j,t+1}) + e_{i,t+1}
\]

(5.3.4)

for \( i = 1, \ldots, N \), where \( s_{i,t} \) stands for the exchange rate of the \( i \)th currency against the \( N+1 \)st and where \( e_t = (e_{1,t}, \ldots, e_{N,t}) \) is i.i.d. normal with mean zero and time-varying variance-covariance matrix \( \mathbf{H}_t \). Alternatively (5.3.4) can be derived as an approximation to more general models of the risk premium.

The second extension of the work by DH85 which BB90 consider consists in replacing the ARCH specification in DH85 by a multivariate GARCH(1, 1) model with constant conditional correlations:

\[
    h_{ii,t} = \psi_i + \beta_i h_{ii,t-1} + \alpha_i e_{i,t-1}^2
\]

(5.3.5)

\[
    h_{ij,t} = \rho_{ij} \sqrt{h_{ii,t}} \sqrt{h_{jj,t}}
\]

if \( i \neq j \)

where \( h_{ij,t} = \{\mathbf{H}_t\}_{ij} \). This specification avoids the implausible lag patterns in DH85 and has the advantage over alternative parametrizations of multivariate GARCH that consistent parameter estimates can be easily obtained under the assumption that the impact of the conditional covariances on the conditional mean can be ignored. If \( \delta_{ij} = 0 \) (\( \forall \ i, j \)) univariate GARCH(1, 1) models can be estimated for all bilateral exchange rates separately and subsequently the conditional correlation parameters \( \rho_{ij} \) can be estimated as the correlations of the rescaled disturbances \( e_{i,t} / \sqrt{h_{ii,t}} \). In this way suitable starting values for the computationally demanding full numerical optimization of the likelihood function can be obtained as long as the GARCH-M effect is not too dominant.

The third extension in BB90 is their use of weekly instead of monthly data. If it is assumed that a month consists of four weeks, the model proposed for weekly data can be written as:

\[
    s_{i,t+4} - f_i^{(4)} = \mu_i + \sum_{j=1}^{N} \delta_{ij} \text{cov}_t(s_{i,t+1}, s_{j,t+1}) + e_{i,t+4}
\]

(5.3.6)

where \( f_i^{(4)} \) is the monthly forward rate in week \( t \) of currency \( i \). As by definition \( e_{i,t+4} = s_{i,t+4} - E_t[s_{i,t+4}] \), \( e_{i,t} \) is generated by a third-order moving average process, because both \( E_{t-k}[s_{i,t}] \) and \( s_{i,t} \) are included in the information set at period \( t \) if \( k \geq 3 \). If weekly exchange rate changes are uncorrelated, which is well known to be approximately true, the first three autocorrelations of \( e_{i,t} \) will be 0.75, 0.50 and 0.25 respectively. In fact a month contains more than four weeks, which yields a MA(4) process and slightly different autocorrelations
(see BB90). BB90 impose orthogonality of the spot rate innovations in the various currencies \((Ee_{t,s}e_{t,s} = 0 \forall t, s, i, j)\) and write:

\[
e_{i,t+4} = e_{i,t+4} + \sum_{k=1}^{4} \vartheta_k e_{i,t+4-k}
\]

where the \(\vartheta_k\) are the moving average coefficients which yield the autocorrelations referred to above. These values of \(\vartheta_k\) are imposed throughout the numerical maximization of the likelihood function corresponding to (5.3.5), (5.3.6) and (5.3.7).

In the empirical section of the paper BB90 restricted themselves to four major European currencies, the pound sterling, the German mark, the Swiss franc and the French franc, all against the US dollar. They consider 462 weekly opening prices from the New York Foreign Exchange Market between March 1, 1980 and February 2, 1989. Imposing \(\delta_{ij} = 0\) initially, the authors compute Ljung–Box tests for autocorrelation in the residuals and find no evidence of autocorrelation in the residuals in addition to the imposed MA(4) process. Remember that, as discussed in section 4.3, the presence of conditional heteroskedasticity generally induces an upward bias in the traditional test statistics for the absence of autocorrelation. Estimation of the model while retaining the restriction that \(\delta_{ij} = 0\) reveals significant conditional heteroskedasticity.

Instead of estimating all parameters in (5.3.5), (5.3.6) and (5.3.7) simultaneously, BB90 restrict themselves to the computation of Lagrange Multiplier tests for the hypothesis \(H_0: \delta_{ij} = 0\). One of the tests carried out by BB90 is a currency-by-currency test of the joint significance of the own conditional variance and the conditional covariances with the other currencies. Again, the results are somewhat disappointing. The four test statistics do not lend much support to the idea that the risk premium is a simple linear function of the corresponding covariances as specified in (5.3.4). Only for the UK is there some evidence that the conditional covariances explain part of the risk premium in addition to the own conditional variances.

5.4. The Monthly Non-Parametric GARCH-M Model for the Risk Premium

The disappointing empirical results in DH85 and BB90 could be due to the fact that the underlying model is an insufficiently accurate description of reality. Several other explanations of the empirical results are possible, however, such as the limited sample size in BB90 and the assumed conditional normality of the disturbances in both papers. In this section we will discuss yet another potential cause of the failure to find empirical results in line with the theory: misspecification of the conditional variance equation, which is based on an auxiliary assumption that is not derived from theory. In section 4.2 we have already referred to the fact that in GARCH-M models incorrect specification
of the variance equation will lead to inconsistent estimates of the mean parameters. Fortunately, however, Pagan and Ullah (1988) have suggested tests for the specification of the conditional variance and have proposed estimation strategies which are more robust to the specification of the variance equation.

Let us start the discussion of Pagan and Ullah (1988) (PU88 from now on) by reconsidering equation (5.2.1) assuming that \( n_t \) is non-stochastic. If \( n_t \) is non-stochastic, (5.2.1) can be written as:

\[
s_{t+1} - f_t = \mu + \delta \text{var}_t[m_{t+1}] + e_{t+1}
\]

where \( e_{t+1} \) is a zero mean, uncorrelated but possibly conditionally heteroskedastic error term. It is important to note that in (5.4.1) the conditional variance of an exogenous variable appears as a regressor. One possible estimation strategy for (5.4.1) starts off by estimating the AR(1) model which generates \( m_t \) by assumption. The resulting error terms will be denoted by \( \tilde{\sigma}_{mt} = m_t - \hat{\rho}_m m_{t-1} \). The second step in this estimation strategy is to replace \( \text{var}_t[m_{t+1}] \) by \( \tilde{\sigma}_{mt+1}^2 \), which yields:

\[
s_{t+1} - f_t = \mu + \delta \tilde{\sigma}_{mt+1}^2 + e_{t+1} + \delta \{ \text{var}_t[m_{t+1}] - \tilde{\sigma}_{mt+1}^2 \}
\]

Ordinary least squares estimation of (5.4.2) will not in general yield consistent estimates of \( \mu \) and \( \delta \) as \( \tilde{\sigma}_{mt+1}^2 \) will in non-degenerate cases be correlated with \( \{ \text{var}_t[m_{t+1}] - \tilde{\sigma}_{mt+1}^2 \} \). While OLS is inconsistent, instrumental variables estimators will yield consistent estimates of \( \mu \) and \( \delta \) if the instruments are observed at time \( t \), are uncorrelated with \( e_{t+1} \) and are correlated with the regressors. Instruments such as a constant term, and present and lagged values of \( m_t^2 \), will satisfy these requirements under very general assumptions. In this way consistent estimators of the impact of \( \text{var}_t[m_{t+1}] \) on \( E_t[s_{t+1} - f_t] \) can be obtained without specifying the functional form of \( \text{var}_t[m_{t+1}] \), i.e. without making a choice between an ARCH, GARCH, EGARCH model or any other functional form.

Instrumental variables estimation of (5.4.2) has the advantage of being robust against potential misspecification of the variance equation, but it can be very inefficient if the variance of \( \delta \{ \text{var}_t[m_{t+1}] - \tilde{\sigma}_{mt+1}^2 \} \) is large compared to that of \( e_{t+1} \). Similar issues of robustness versus efficiency arise in the literature on the estimation of linear models with unobserved rational expectations; see e.g. Nijman (1990) and Nijman and Palm (1991). This literature suggests that more efficient robust IV estimators can be derived by considering better proxies for \( \text{var}_t[m_{t+1}] \), i.e. by adopting the substitution approach to replace unobserved variables instead of the errors in variables approach. One possible proxy for \( \text{var}[m_{t+1} | I_t] \) is \( \text{var}[m_{t+1} | H_t] \) where \( H_t \subset I_t \) and \( I_t \) is the set of all
information which is available at time $t$. Note that $\text{var}[m_{t+1}] = \text{var}[m_{t+1} \mid I_t]$ by definition. Substitution of $\text{var}[m_{t+1} \mid H_t]$ in (5.4.1) yields:

$$s_{t+1} - f_t = \mu + \delta \text{var}[m_{t+1} \mid H_t] + u_{t+1}$$

(5.4.3)

where:

$$u_{t+1} = e_{t+1} + \delta(\text{var}[m_{t+1} \mid I_t] - \text{var}[m_{t+1} \mid H_t])$$

(5.4.4)

These equations show that a regression of the forward rate forecast errors on the variance of the money supply conditional on a subset of all information will yield consistent estimates of $\mu$ and $\delta$. Non-parametric estimates of this variance which do not depend on arbitrary assumptions on the functional form can be obtained, e.g. using kernel estimators of conditional means by noting that:

$$\text{var}[m_{t+1} \mid H_t] = E[m_{t+1}^2 \mid H_t] - (E[m_{t+1} \mid H_t])^2$$

(5.4.5)

The kernel estimator of a conditional mean with a finite number of conditioning variables reads as:

$$\hat{E}[y_t \mid x_t] = \frac{1}{T} \sum_{s=1}^{T} y_s K(\gamma T^{-1} (x_t - x_s)) \left( \sum_{s=1}^{T} K(\gamma T^{-1} (x_t - x_s)) \right)$$

(5.4.6)

where $K(\cdot)$ is a kernel function that aims to smooth the data and where $\gamma_T$ is the so-called bandwidth parameter that is typically proportional to $T^{-1/4+q}$ and $q$ is the dimension of the conditioning set. Many types of kernel might be employed. A popular one is the multivariate normal kernel. Under various restrictions on the bandwidth parameter $\gamma_T$ and assumptions on the processes generating $(y_t, x_t)$ it has been shown that $\hat{E}[y_t \mid x_t]$ converges to $E[y_t \mid x_t]$ in probability. Recent applications of kernel estimation to modelling conditional heteroskedasticity include Pagan and Hong (1990) and Sentana and Wadhwani (1989). A good illustration and comparison of different parametric and non-parametric methods of modelling conditional variances is given in Pagan and Schwert (1990).

The empirical results in PU88 still show little sign of impact of the conditional variance on the conditional mean. One explanation is that their empirical results are obtained from monthly data, which do not show much conditional heteroskedasticity. Obviously, another one is that the underlying model does not hold.

5.5. Concluding Remarks on Risk Premia in Foreign Exchange

In this section we have discussed the estimation of structural econometric models of the risk premium. The models that we considered suggest that the
risk premium depends on conditional variances and possibly conditional covariances, either of money supplies, or of spot rates. However, the empirical results in neither of the three papers that we discussed, nor in other papers in this area, support the overly simple models which they tested. More sophisticated models to explain time-varying risk premia are required. In particular models in which the risk premium is a function of the conditional variances of a set of exogenous variables seem promising. Hodrick (1989) provides an example of such a model, showing how the exchange rate is affected by uncertainty in the monetary policy, government expenditures and income growth. Implementation and tests of these models can of course be naturally conducted within the framework described in this chapter.

6. Models of Time-Varying Premia in Stock or Bond Returns

6.1. Introduction

In section 2.2 we introduced a capital asset pricing model (CAPM) which can be used to take portfolio decisions in a world where returns have time-varying first and second conditional moments. In order to test the model, or to use the model to achieve an optimal trade-off between expected returns and unhedged risks, estimates of these time-varying moments are required. In this section several multivariate GARCH models will be considered which can be used to model first and second conditional moments of a vector of returns.

In the model presented in Section 2.2, the conditional mean returns depend on the conditional covariances of the returns as in (2.2.5), which we repeat here for convenience:

$$E_t[r_{t+1}] - r^f_t = \rho_t \text{ var}_t[r_{t+1}] x_t$$

(6.1.1)

where $r_{t+1}$ is an $N \times 1$ vector of returns in period $t+1$ on the $N$ assets in the economy, $r^f_t$ is the risk-free rate in period $t+1$ which is known in period $t$, $t$ is an $N$-dimensional vector of ones, $x_t$ is an $N \times 1$ vector of investment shares and $\rho_t$ is the price of risk, which will for simplicity be assumed to be time-invariant: $\rho_t = \rho$. Equation (6.1.1) can be derived under much more general assumptions, as in Campbell (1990).

Premultiplication of (6.1.1) with $x_t$ yields a univariate relation between the conditional mean and variance of the market portfolio. Several authors (see e.g. French, Schwert and Stambaugh (1987)) have used univariate GARCH-M models to describe excess returns on the market portfolio, and found a significant impact of the conditional variance on the conditional mean in line with the underlying CAPM. We shall test (6.1.1) using multivariate GARCH models. Note also that multivariate GARCH models can be used to derive hedging strategies whether or not (6.1.1) holds, as the optimal portfolio for
investors who maximize (2.2.1) will still be given by (2.2.4). Obviously the restriction which was imposed in section 2.2 that the economy consists of representative investors can easily be relaxed by allowing for differences in risk aversion.

In section 6.2 we shall first of all consider the diagonal multivariate GARCH model which Bollerslev, Engle and Wooldridge (1988) (BEW88 from now on) used to describe the returns on the US bills, bonds and stock market. As the number of parameters in unrestricted variance equations of multivariate GARCH models soon gets unmanageable, several authors have tried to find parametrizations which impose a priori plausible restrictions. In section 6.3 we shall consider one model in this line of research, the FACTOR-ARCH model which is used by Engle, Ng and Rothschild (1990) to model the term structure of interest rates. Section 6.4 concludes.


The conceptually most straightforward generalization of the analysis of the relation between the conditional mean and variance of the returns on the market portfolio to the multivariate case is to consider the multivariate GARCH\((P, Q)\)-M model

\[
\begin{align*}
  r_{t+1} - r_t^f &= \mu + \rho \, \text{var}_t[r_{t+1}] \, x_t + \varepsilon_{t+1} \\
  \mathbb{E}[\varepsilon_{t+1}^2 | I_t] &= \text{var}_t[r_{t+1}] = \Sigma_t
\end{align*}
\]  

(6.2.1)

\[
\sigma_i(i, j) = \psi_{ij} + \sum_{r,q=1}^{N} \sum_{k=1}^{P} \beta_k(i, j, r, q) \sigma_{t-k}(r, q) + \sum_{r,q=1}^{N} \sum_{k=1}^{Q} \alpha_k(i, j, r, q) \varepsilon_{t-k} e_{r,q,t-k}
\]  

(6.2.3)

where \(\varepsilon_{t+1} = r_{t+1} - E_t[r_{t+1}]\) and where \(\sigma_i(i, j)\) denotes the \((i, j)\)th element of \(\Sigma_t\). Sufficient conditions on the \(\alpha\)s and \(\beta\)s which ensure that \(\Sigma_t\) will be positive definite have been given in Baba et al. (1989). As \(x_t\) is observable, the system in (6.2.1), (6.2.2) and (6.2.3) can be estimated in principle from data on \(r_{t+1}\), \(r_t^f\) and \(x_t\) for \(t = 1, \ldots, T\). One of the implications of (6.1.1) is that \(\mu = 0\), which can be tested. In practice, estimation of the unrestricted model in (6.2.1) to (6.2.3) is impossible, however, unless \(N\) is very small, as the number of parameters tends to be extremely large. The symmetry of \(\Sigma\) implies that the number of parameters in the variance equation is \(0.5N(N + 1) + 0.25N^2(N + 1)^2(P + Q)\) which if \(P = Q = 1\) yields a value of 12 if \(N = 2\), of 42 if \(N = 3\), of 110 if \(N = 4\), etc. Obviously restrictions are required to keep the estimation problem manageable.
A drastic simplification of (6.2.3) is obtained if one is willing to assume that \( P = Q = 1 \) and that \( \alpha(i, j, q, r) = 0 \) and \( \beta(i, j, q, r) = 0 \) unless \( i = q \) and \( j = r \), in which case \( \alpha(i, j, q, r) = \alpha_{ij} \) and \( \beta(i, j, q, r) = \beta_{ij} \). The model which is obtained if these restrictions are imposed, as in BEW88, is known as the diagonal model. If \( P = Q = 1 \) the diagonal equivalent of (6.2.3) can be written as:

\[
\sigma_t(i, j) = \psi_{ij} + \beta_{ij} \sigma_{t-1}(i, j) + \alpha_{ij} \epsilon_{i,t-1} \epsilon_{j,t-1} \tag{6.2.4}
\]

which contains \( 1.5 N (N + 1) \) parameters. If \( N = 3 \), as in the application in BEW88, the number of parameters in (6.2.4) is 18, as opposed to 42 in (6.2.3). Drawbacks of the diagonal model are that the restrictions are quite arbitrary and that the conditional covariance matrix generated by (6.2.4) is not necessarily positive semi-definite.

In the empirical section of their paper BEW88 consider quarterly returns on six-month Treasury bills, twenty-year Treasury bonds and stocks from 1959-I to 1984-II. The return on three-month Treasury bills is taken to represent the risk-free return. The plots of the three excess holding yields which are given in BEW88 clearly suggest that not only their conditional means but also the conditional variances vary over time, in line with (6.1.1).

The model which is estimated in BEW88 consists of (6.2.1), (6.2.2) and (6.2.4). The estimation results suggest a reasonable and significant estimate of the price of risk parameter \( \rho \) as well as a time-varying variance for excess returns on six-month bills, a slightly time-varying variance for excess returns on government bonds, but no time-varying variance for excess returns on stocks. All three intercepts \( \mu_t \) which should be insignificant if (6.1.1) holds, are in fact significant. BEW88 explain the negative intercept for bonds and stocks from the fact that capital gains are not as heavily taxed as dividend and interest payments. This provides incentives to hold these assets even at otherwise unfavourable rates of return.

BEW88 test the validity of the CAPM relation in (6.1.1) using Lagrange Multiplier tests for the inclusion of additional regressors in the mean equation (6.2.1). The test for the inclusion of the own conditional variances (6.2.2) as regressors in addition to the intercept and the covariance with the market portfolio does not reject (6.1.1). That is of particular interest, since in tests of the time-invariant CAPM the own variance is often found to be highly significant. Tests for the inclusion in (6.2.1) of lagged excess holding yields on the one hand, or innovations in the logarithm of per capita consumption on the other hand, reject the model in (6.2.1), (6.2.2) and (6.2.4) very clearly. One reason might of course be that (6.1.1) does not hold, but another reason could be that the variance equation (6.2.4), which is not derived from theory, is misspecified. Misspecification of (6.2.4) arises if the restrictions imposed by the diagonal model are not satisfied, but another explanation might be that premia and conditional heteroskedasticity depend on information in addition to past innovations in asset returns. In particular the Lagrange Multiplier tests for omitted variables in (6.2.1) suggest that lagged excess holding yields and
innovations in consumption might have some explanatory power when added to (6.2.4). Giovannini and Jorion (1989) have extended the diagonal multivariate GARCH model by including cross products $d_{it-1}d_{jt-1}$ in (6.2.4), where $d_i$ is the difference in returns on government debt in country $i$ with the return on government debt in the US.

### 6.3. The Factor-ARCH Model for the Term Structure of Interest Rates Proposed by Engle, Ng and Rothschild (1990)

From section 6.2 it is evident that the most important difficulty in the extension of univariate models for time-varying risk premia to multivariate models, is the fact that in unrestricted models the number of parameters soon becomes unmanageable. In a recent paper, Engle, Ng and Rothschild (1990) (ENR90 from now on) have proposed FACTOR-ARCH models as a parsimonious structure for the conditional covariance matrix of asset excess returns. FACTOR-ARCH models are appealing because they model the notion that the risk on financial markets can be decomposed into a limited number of factors and an asset specific ('idiosyncratic') error term. A similar model arises from the Arbitrage Pricing Theory (APT) although the APT does not imply that the number of factors is finite. The FACTOR-ARCH model is used to model interest rate risk in ENR90, while a companion paper (Ng, Engle and Rothschild (1992)) considers risk premia and anomalies to the CAPM on the US stock market.

One way to generate the model which is used in ENR90 is to assume that the vector of excess returns, $r_{t+1} - r^f$, is generated by a factor structure in which the factors are conditionally heteroskedastic:

$$r_{t+1} - r^f_t = \mu_t + \sum_{k=1}^{K} \beta_k f_{kt} + \nu_t$$

(6.3.1)

where $r_{t+1}$ is an $N$-dimensional vector and $K < N$ is the number of factors. In (6.3.1) $\mu_t$ is an $N$-dimensional vector of risk premia to be determined and $\beta_k$ is an $N$-dimensional vector of coefficients. It is assumed that all factors are independent and have zero mean, that the idiosyncratic errors have zero mean, are potentially correlated but are conditionally homoskedastic. The factors on the other hand are allowed to be conditionally heteroskedastic. Stated formally, the assumptions are:

$$E_{t-1}[\nu_{kt}] = 0, \quad E_{t-1}[f_{kt}] = 0,$$

$$\text{var}_{t-1}[\nu_{kt}] = \Omega_{kk}, \quad \text{var}_{t-1}[f_{kt}] = \lambda_{kt}$$

(6.3.2)

$$\text{cov}_{t-1}[\nu_{kt}, \nu_{rt}] = \Omega_{kr}, \quad \text{cov}_{t-1}[f_{kt}, f_{rt}] = 0 \quad (k \neq r)$$

ENR90 assume that the return on the market portfolio satisfies (6.3.1) and
(6.3.2) as well, which is the case if the optimal portfolio weights $x_t$ are not time-varying ($x_t = x \ (\forall \ t)$). The risk premium on the market portfolio is referred to as $\mu^m_t$, while the ‘beta’ of the market with respect to the $k$th factor is referred to as $\beta^m_k$, i.e. as a matter of notation one can write:

$$r^m_{t+1} - r^f_t = \mu^m_t + \sum_{k=1}^K \beta^m_k f_{kt} + \nu^m_t$$  \hspace{1cm} (6.3.3)

The assumption on the covariance structure of the asset returns in (6.3.1), (6.3.2) and (6.3.3) can be used to derive an expression for the risk premia in the individual assets in which the portfolio weights no longer occur.

Assuming once more that the price of risk is time-invariant, along the lines of section 2.2, expression (6.1.1) can be written as:

$$E_t[r_{t+1}] - r^f_t = \rho \text{cov}_t[r_{t+1}, r^m_{t+1}]$$  \hspace{1cm} (6.3.4)

Substituting (6.3.1) and (6.3.3), and assuming that the idiosyncratic disturbance on the market return is negligible ($\text{var}_t[\nu^m_{t+1}] = 0$) one obtains

$$E_t[r_{t+1}] - r^f_t = \rho \sum_{k=1}^K \beta^m_k \text{var}_t[f_{kt+1}] \beta_k$$  \hspace{1cm} (6.3.5)

which yields a direct expression for the risk premium.

As the factors $f_k$ themselves are not observable, ENR90 subsequently define the concept of a factor-representing portfolio. The portfolio with portfolio weights $\alpha_k = (\alpha_{k1}, \ldots, \alpha_{kN})$ is referred to as a factor-representing portfolio for factor $k$ if its return is uncorrelated with all factors except factor $k$, and if the (conditional) covariance with factor $k$ coincides with the (conditional) variance of factor $k$, i.e.:

$$\text{cov}_t[\alpha_k r_{t+1}, f_{kt+1}] = \text{var}_t[f_{kt+1}] = \lambda_{kt+1}$$  \hspace{1cm} (6.3.6)

$$\text{cov}_t[\alpha_k r_{t+1}, f_{qt+1}] = 0 \hspace{1cm} (k \neq q)$$

from which it follows that $\alpha_k^t \beta_k = 1$ and $\alpha_k^t \beta_l = 0 \ (k \neq l)$. ENR90 subsequently show that the risk premium on any asset can be expressed as a linear combination of the premia on the factor representing portfolios, and therefore can be expressed as a linear function of the conditional variances of the factor representing portfolios only

$$E_t[r_{t+1}] - r^f_t = \sum_{k=1}^K \{E_t[\alpha_k r_{t+1}] - r^f_t\} \beta_k$$

$$= \sum_{k=1}^K \{\gamma_k \text{var}_t[\alpha_k r_{t+1}] + c_k\} \beta_k = \mu_t$$  \hspace{1cm} (6.3.7)
Comparing (6.1.1) with (6.3.7) it is apparent that the main achievement of the FACTOR-ARCH model is that the risk premium is expressed as a function of the conditional variances of $K$ portfolios rather than as a function of the conditional covariance matrix of all $N$ returns.

In order to complete the model, the factor-representing portfolios and a specification for $\text{var}_t[\alpha_k r_{t+1}]$ ($k = 1, \ldots, K$) have to be chosen. Although it is in principle possible to choose the factor-representing portfolios by estimating the weights $\alpha_k$ and all other parameters in the model jointly, using a numerical maximization of the likelihood over a large number of parameters, this approach is not taken in ENR90. Instead, a priori knowledge of the number of factor-representing portfolios and their weights is assumed, which leaves the specification of $\text{var}_t[\alpha_k r_{t+1}]$ as the final point to consider.

The simplest but most restrictive assumption on the dynamics of the conditional variances of the returns on the factor-representing portfolios is that these returns are not only uncorrelated, $\text{cov}_t[\alpha_k r_{t+1}, \alpha_q r_{t+1}] = 0$ ($k \neq q$), but that moreover shocks in one factor do not affect the conditional variance of the other factors in any way, i.e.:

$$\text{var}_t[\alpha_k r_{t+1} | r_t, r_{t-1}, \ldots] = \text{var}_t[\alpha_k r_{t+1} | \alpha_k r_t, \alpha_k r_{t-1}, \ldots]$$ (6.3.8)

which is referred to by ENR90 as a ‘univariate portfolio representation’.

If one is willing to assume that the factor-representing portfolios have a univariate portfolio representation and that the conditional variances of the returns on the factor-generating portfolios are generated by GARCH(1,1) models, the risk premia on the factor-representing portfolios can be derived by estimating the univariate GARCH-M model:

$$\alpha_k r_{t+1} - r_t^f = c_k + \gamma_k \delta_{k,t+1} + e_{k,t+1}$$ (6.3.9)

$$\delta_{k,t+1} \sim N(0, \sigma_{k,t+1})$$

$$\delta_{k,t+1} = \psi_k + \varphi_k \delta_{k,t} + \omega_k (e_{k,t})^2$$

where we have changed the standard notation of the variance equation to avoid confusion with the $\beta_k$ and $\alpha_k$ expressions in (6.3.1) to (6.3.7). Adding a normality assumption to (6.3.7) one easily obtains:

$$r_{i,t+1} - r_t^f = \sum_{k=1}^{K} \beta_{ik} (c_k + \gamma_k \delta_{k,t+1}) + e_{i,t+1}$$ (6.3.10)

$$e_{i,t+1} \sim N\left(0, \sigma_i + \sum_{k=1}^{K} \beta_{ik}^2 \delta_{k,t+1}\right)$$

A simple way to obtain consistent estimates of the $\beta_{ik}$ therefore is to maximize the likelihood of the $r_{i,t+1}$ over $\sigma_i$ and $\beta_{ik}$ treating the $\delta_k$, $\gamma_k$ and $\delta_{k,t+1}$
expressions as if they coincided with the true values of these parameters. In this way only a small number of relatively simple numerical optimization problems as implied by (6.3.9) and (6.3.10) have to be solved to estimate the full model.

In their empirical application, ENR90 use data on monthly returns on US Treasury bills and on the value-weighted index of NYSE and AMEX stocks from August 1964 to November 1985. The one-month T-bill rate is taken as the riskless return, and the FACTOR-ARCH model is used to model the excess returns on the two- to twelve-month T-bills using an equally weighted portfolio of all T-bills and the stock market portfolio as the \emph{a priori} chosen factor representing portfolios (i.e. \(K = 2\)). The first portfolio is referred to as EWB (equally-weighted bills), the second one as VWS (value-weighted stocks). The estimates in ENR90 unambiguously reveal that the risk premia and volatilities of T-bills with longer maturities are more sensitive to changes in the conditional variance of the first factor representing portfolio. The \(\beta\)s with respect to the VWS-portfolio tend to be insignificant, suggesting that the data-generating process might in fact be a one-factor model with EWB as the only factor. A battery of tests is used to test the model specification and generally yields results which are supporting for the model.

6.4. Concluding Remarks

Several multivariate GARCH models which can be used to model first- and second-order conditional moments have been discussed in this section. The state of the art model appears to be the FACTOR-ARCH model considered in section 6.3. This model appears to be a powerful but manageable tool. While the models in section 6.2 require high dimensional numerical optimization, the FACTOR-ARCH model can be analysed by estimating a number of relatively simple models only. Note, however, that the important problems of the choice of the number of factors and the weights of the factor-representing portfolios are not addressed by ENR90. Some hints toward possible solutions are given in the companion paper, Ng, Engle, and Rothschild (1992), which contains an application on testing for a firm size anomaly in a conditional CAPM for US stock returns.

7. Conclusions

In this chapter, models for time-varying volatility measures have been presented and their application to and relevance for the analysis of financial series have been illustrated using examples from the finance literature. In a way, GARCH models are natural extensions of ARMA schemes to describe the time dependencies in second moments of many economic series. These non-linear models are fairly easily implemented, estimated and tested. The standard apparatus of autocorrelation function analysis can be applied to the
series squared and used to empirically determine the order of the GARCH process. Also, these models have been found to be consistent with many stylized facts of financial series, such as fat-tailed marginal distributions and zero serial autocorrelation of a series but dependencies over time characterized by tranquil and volatile periods. Further extensions to non-parametric methods and other functional form specifications for conditional second moments of economic time series are on the research agenda in this area. More applications of the models are expected to contribute to a better understanding of the time series properties of financial and other economic series.

References


