Autoregressive Conditional Heteroskedasticity (ARCH) Models: A Review

Degiannakis, Stavros and Xekalaki, Evdokia

Department of Statistics, Athens University of Economics and Business

2004

Autoregressive Conditional Heteroscedasticity (ARCH) models:A Review

Stavros Degiannakis and Evdokia Xekalaki
Department of Statistics, Athens University of Economics and Business, Greece

Abstract

Autoregressive Conditional Heteroscedasticity (ARCH) models have successfully been employed in order to predict asset return volatility. Predicting volatility is of great importance in pricing financial derivatives, selecting portfolios, measuring and managing investment risk more accurately. In this paper, a number of univariate and multivariate ARCH models, their estimating methods and the characteristics of financial time series, which are captured by volatility models, are presented. The number of possible conditional volatility formulations is vast. Therefore, a systematic presentation of the models that have been considered in the ARCH literature can be useful in guiding one's choice of a model for exploiting future volatility, with applications in financial markets.

Keywords and Phrases: ARCH models, Forecast Volatility.

1. Introduction

Since the first decades of the 20th century, asset returns have been assumed to form an independently and identically distributed (i.i.d) random process with zero mean and constant variance. Bachelier (1900) was the first who contributed the theoretical random walk model for the analysis of speculative prices. For $\{P_t\}$ denoting the discrete time asset price process and $\{y_t\}$ denoting the process of the continuously compounded returns, defined by $y_t = \log(P_t / P_{t-1})$, the early literature viewed the system that generates the asset price process as a fully unpredictable random walk process:

$$P_t = P_{t-1} + \varepsilon_t$$
$$\varepsilon_t \overset{i.i.d.}{\sim} N(0, \sigma^2),$$

where ε_t is a zero-mean i.i.d. normal process. However, the assumptions of normality, independence and homoscedasticity do not always hold with real data.

Figures 1 to 3 depict the continuously compounded daily returns of the Chicago Standard and Poor's 500 Composite (S&P500) index, Frankfurt DAX30 stock index and Athens Stock Exchange (ASE) index. The data cover the period from 2nd January 1990 to 27th June 2000. A visual inspection shows clearly, that the mean is constant, but the variance changes over time, so the return series is not a sequence of independently and identically distributed (i.i.d.) random variables. A characteristic of asset returns, which is noticeable from the figures, is the volatility clustering first noted by Mandelbrot (1963): "Large changes tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes". Fama (1970) also observed the alternation between periods of high and low volatility: "Large price changes are followed by large price changes, but of unpredictable sign".

A non-constant variance of asset returns should lead to a non-normal distribution. Figure 4 represents the histograms and the descriptive statistics of the stock market series plotted in Figures 1 to 3. Asset returns are highly leptokurtic and slightly asymmetric, a phenomenon correctly observed by Mandelbrot (1963): "The empirical distributions of price changes are usually too "peaked" to be relative to samples from Gaussian populations … the histograms of price changes are indeed unimodal and their

Figure 1. S&P500 Continuously Compounded Daily Returns from 2/1/90 to 27/06/00

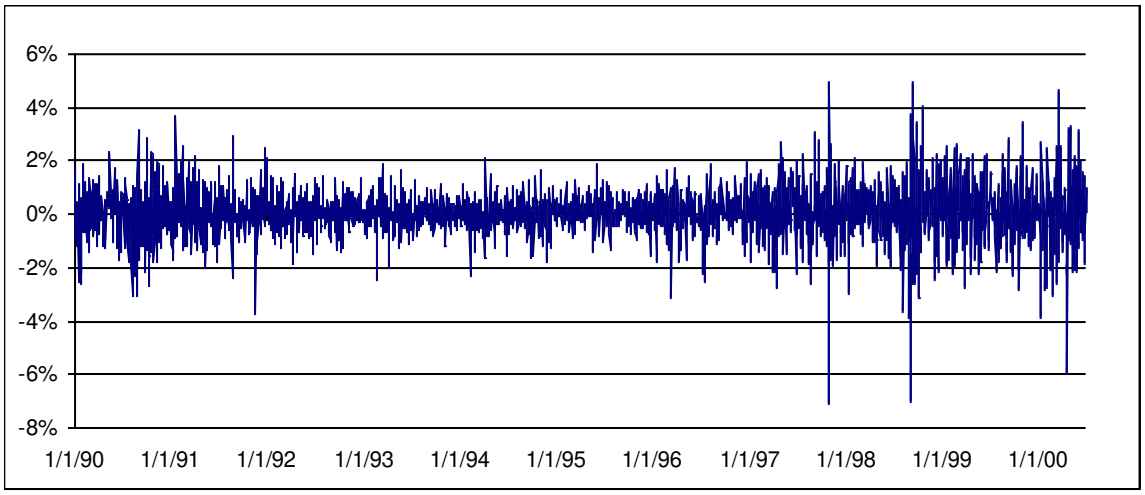


Figure 2. DAX 30 Continuously Compounded Daily Returns from 2/1/90 to 27/06/00

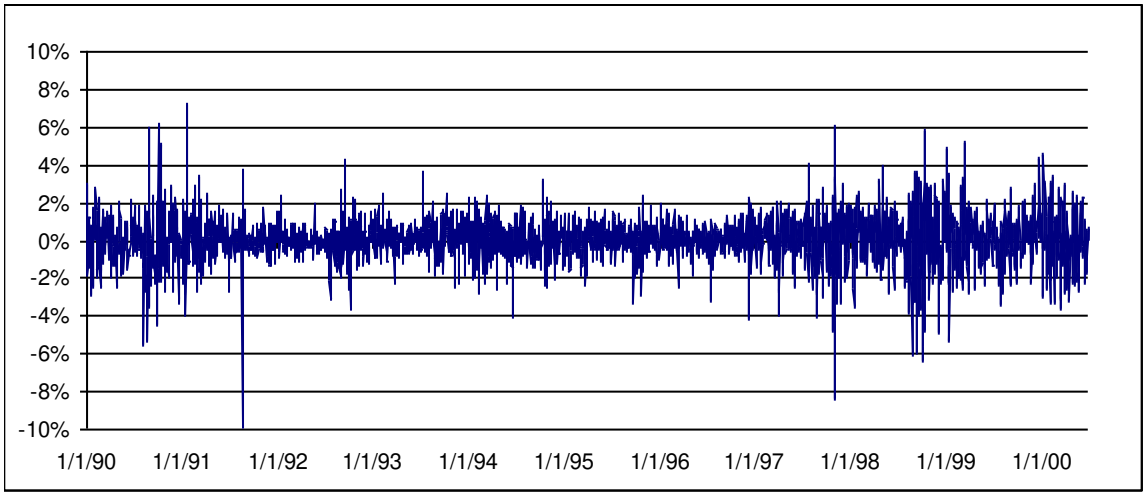
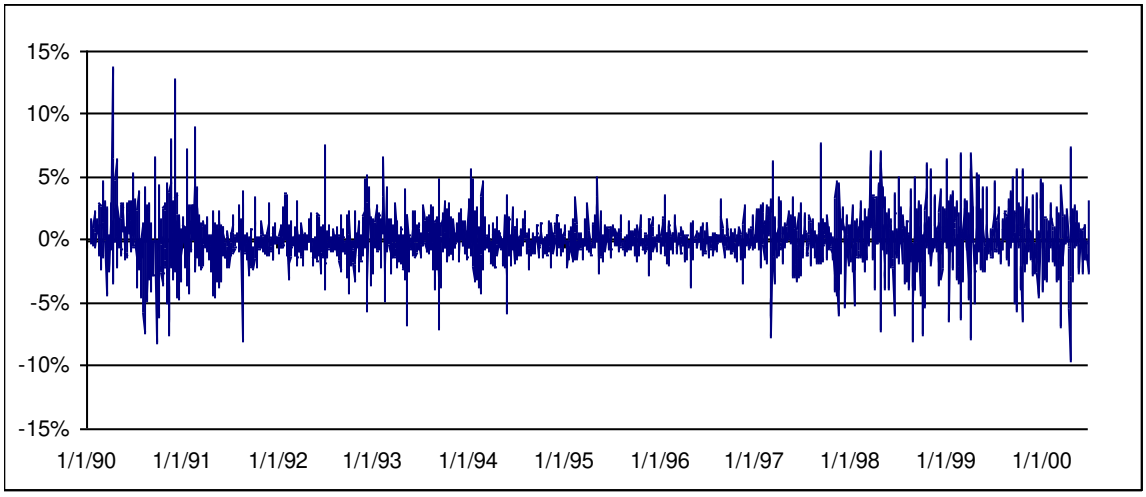
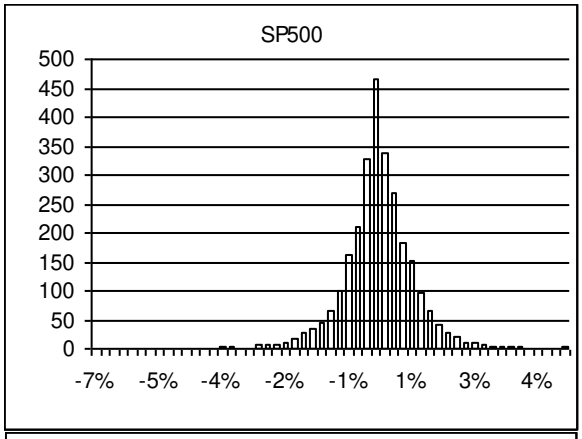


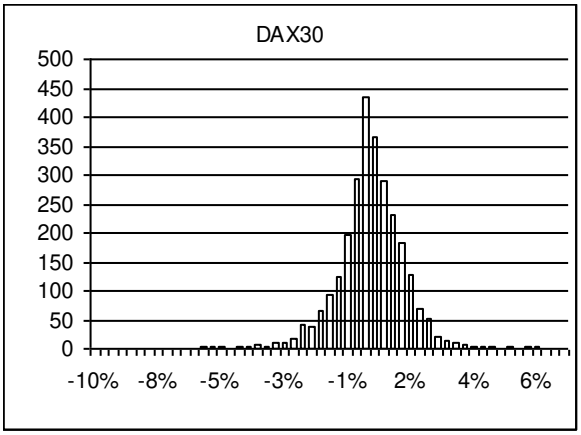
Figure 3. ASE Continuously Compounded Daily Returns from 18/1/90 to 27/06/00

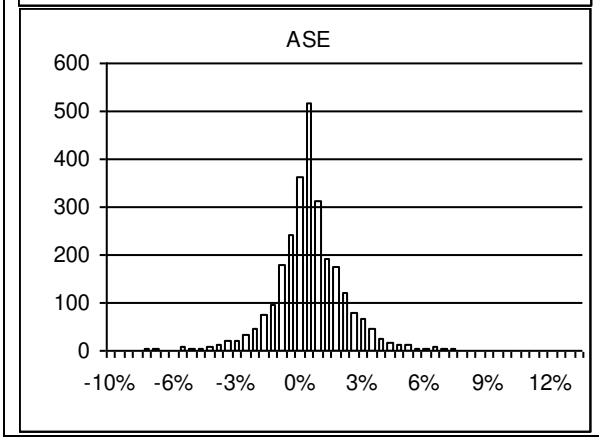


central bells remind the Gaussian ogive. But, there are typically so many outliers that ogives fitted to the mean square of price changes are much lower and flatter than the distribution of the data themselves." In the sixties and seventies, the regularity of leptokurtosis led to a literature on modeling asset returns as independently and identically distributed random variables having some thick-tailed distribution (Blattberg and Gonedes (1974), Clark (1973), Hagerman (1978), Mandelbrot (1963,1964), Officer (1972), Praetz (1972)).

Figure 4. Histogram and Descriptive Statistics for S&P500, DAX 30 and ASE Stock Market Returns.







	S&P500	DAX 30	ASE
Mean	0.05%	0.05%	0.08%
Standard Deviation	0.93%	1.28%	1.91%
Skewness	-0.346	-0.438	0.142
Kurtosis	8.184	7.716	7.349

These models, although able to capture the leptokurtosis, could not account for the existence of non-linear temporal dependence as the volatility clustering observed from the data. For example, applying an autoregressive model to remove the linear dependence from an asset returns series and testing the residuals for a higher-order dependence using the Brock, Dechert and Scheinkman (BDS) test (Brock et al. (1987), Brock et al. (1991), Brock et al. (1996)), the null hypothesis, that the residuals are i.i.d., is rejected.

In this paper, a number of univariate and multivariate ARCH models are presented and their estimation is discussed. The main features of what seem to be most widely used ARCH models are described with emphasis on their practical relevance. It is not an attempt to cover the whole of the literature on the technical details of the models, which is very extensive. (A comprehensive survey of the most important theoretical developments in ARCH type modeling covering the period up to 1993 was given by Bollerslev et al. (1994)). The aim is to give the broad framework of the most important models used today in the economic applications. A careful selection of references is provided so that more detailed examination of particular topics can be made by the interested reader. In particular, an anthology of representations of ARCH models that have been considered in the literature is provided (section 2), including representations that have been proposed for accounting for relationships between the conditional mean and the conditional variance (section 3) and methods of estimation of their parameters (section 4). Generalizations of these models suggested in the literature in multivariate contexts are also discussed (section 5). Section 6 gives a brief description of other methods of estimating volatility. Finally, section 7 is concerned with interpretation and implementation issues of ARCH models in financial applications.

The remaining of the present section looks at the influence that various factors have on a time series and in particular at effects, which as reflected in the data, are known as the “leverage effect”, the “non-trading period effect”, and the “non-synchronous trading effect”.

1.1 The Leverage Effect

Black (1976) first noted that often, changes in stock returns display a tendency to be negatively correlated with changes in returns volatility, i.e., volatility tends to rise in response to “bad news” and to fall in response to “good news”. This phenomenon is termed the “leverage effect” and can only be partially interpreted by fixed costs such as financial and operating leverage (see, e.g. Black (1976) and Christie (1982). The asymmetry present in the volatility of stock returns is too large to be fully explained by leverage effect.

We can observe the phenomenon of “leverage effect” by plotting the market prices and their volatility. As a naïve estimate of volatility at day t, the standard deviation

Figure 5. Daily Log-values and Recursive Standard Deviation of Returns for the S&P500 Stock Market.

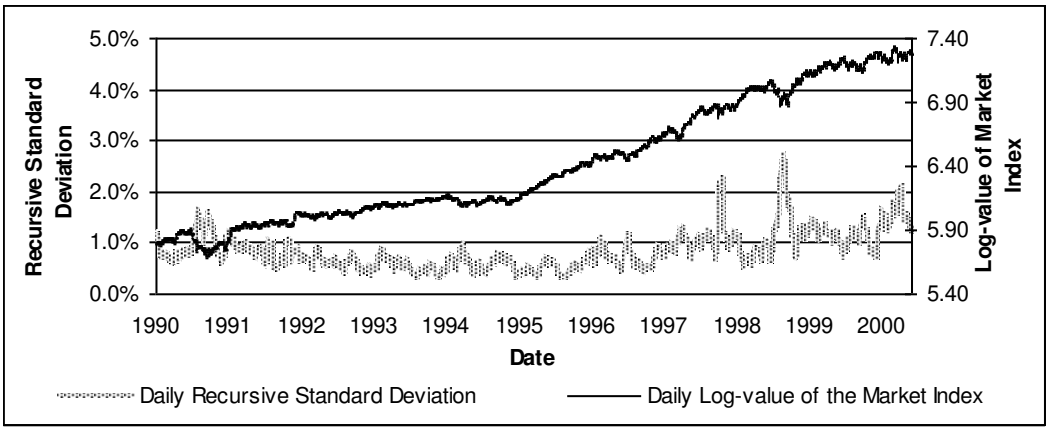


Figure 6. Daily Log-values and Recursive Standard Deviation of Returns for the DAX 30 Stock Market.

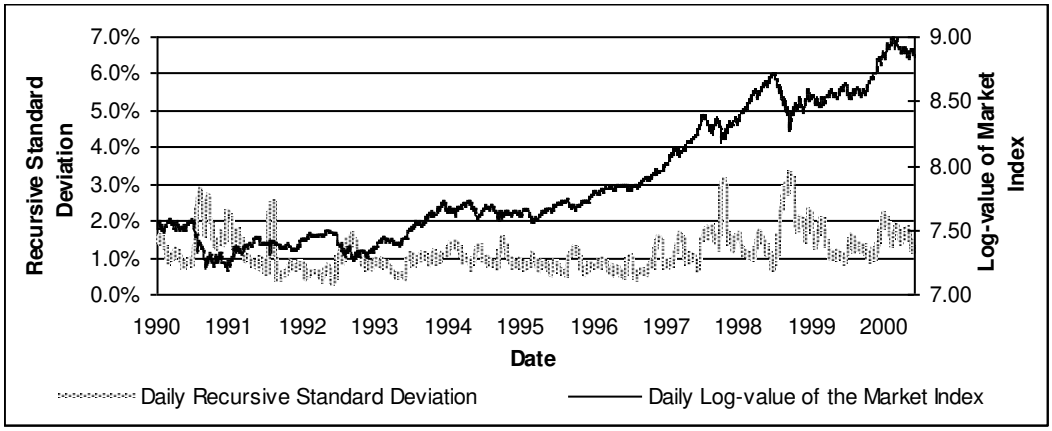


Figure 7. Daily Log-values and Recursive Standard Deviation of Returns for the ASE Stock Market.

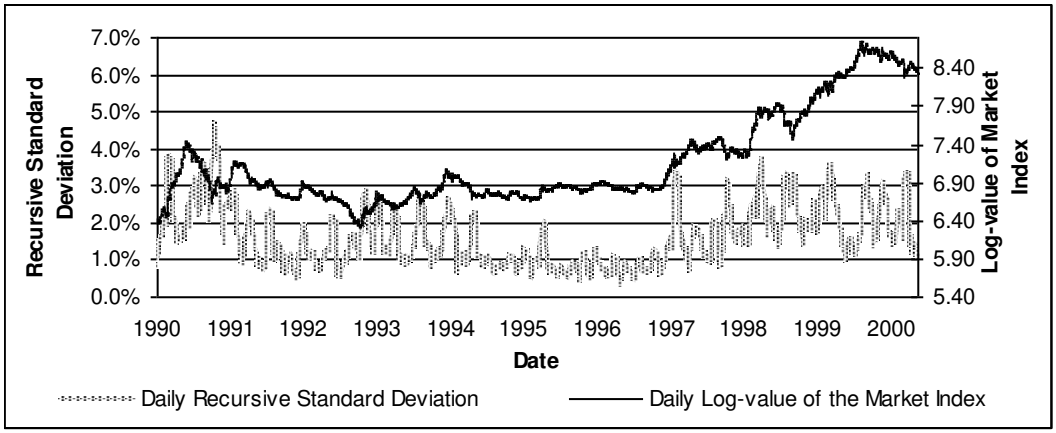


Table 1. Mean and Annualized Standard Deviation of the S&P500, DAX 30 and ASE Index Returns.

	Overall	Monday	Tuesday	Wednesday	Thursday	Friday
			S&P500			
Mean	0.05%	0.12%	0.06%	0.07%	-0.01%	0.04%
St. Deviation	14.80%	15.84%	15.43%	12.57%	14.81%	15.22%
N. of observations	2649	505	543	541	532	528
			DAX 30			
Mean	0.05%	0.07%	0.04%	0.09%	0.00%	0.06%
St. Deviation	20.34%	23.91%	19.79%	18.74%	19.49%	19.46%
N. of observations	2625	518	537	530	516	524
			ASE 500			
Mean	0.08%	0.12%	-0.01%	0.06%	-0.01%	0.26%
St. Deviation	30.27%	39.06%	30.60%	25.98%	28.68%	25.16%
N. of observations	2548	494	523	517	519	495

Annualized standard deviation is computed by multiplying the standard deviation of daily returns by $252^{1/2}$, the square root of the number of trading days per year.

of the 22 most recent trading days, $\sigma_t^{(22)} = \sqrt{\sum_{i=t-22}^{t} \left(y_i - \left(\sum_{i=t-22}^{t} y_i / 22 \right) \right)^2 / 22}$, is used. Figures 5 to 7 plot daily log-values of stock market indices and the relevant standard deviations of the continuously compounded returns. The periods of market drops are characterized by a high increase in volatility.

1.2 The Non-trading Period Effect

Financial markets appear to be affected by the accumulation of information during non-trading periods as reflected in the prices when the markets reopen following a close. As a result, the variance of returns displays a tendency to increase. This is known as the “non-trading period effect”. It is worth noting that the increase in the variance of returns is not nearly proportional to the market close duration as would be anticipated if the information accumulation rate were constant over time. In fact, as Fama (1965) and French and Roll (1986) observed, information accumulates at a lower rate when markets are closed than when they are open. Also, as reflected by the findings of French and Roll (1986) and Baillie and Bollerslev (1989), the returns variance tends to be higher following weekends and holidays than on other days, but not by as much as it would be under a constant news arrival rate. Table 1 shows the annualized standard deviations of stock market returns for each day for the indices S&P500, DAX30 and ASE. The standard deviation on Monday is higher than on other days, mainly for the DAX 30 and ASE indices.

1.3 Non-synchronous Trading Effect

The fact that the values of time series are often taken to have been recorded at time intervals of one length when in fact they were recorded at time intervals of other, not necessarily regular, length is an important factor affecting the return series with an effect known as the "non-synchronous trading effect" (see, e.g. Campbell et al. (1997)). For example, the daily prices of securities, usually analyzed, are the closing prices. The closing price of a security is the price at which the last transaction occurred. The last transaction of each security is not implemented at the same time each day. So, it is falsely assumed that the daily prices are equally spaced at 24-hour intervals. The importance of non-synchronous trading was first recognized by Fisher (1966) and further developed by many researchers such as Atchison et al. (1987), Cohen et al. (1978), Cohen et al. (1979, 1983), Dimson (1979), Lo and MacKinlay (1988, 1990a, 1990b), Scholes and Williams (1977).

Non-synchronous trading in the stocks making up an index induces autocorrelation in the return series, primarily when high frequency data are used. To control this, Scholes and Williams (1977) suggested a first order moving average $\left[MA(1)\right]$ form for index returns, while Lo and MacKinlay (1988) suggested a first order autoregressive $\left[AR(1)\right]$ form. Nelson (1991) wrote "as a practical matter, there is little difference between an $AR(1)$ and an $MA(1)$ when the AR and MA coefficients are small and the autocorrelations at lag one are equal, since the higher-order autocorrelations die out very quickly in the AR model".

2. The ARCH Process

Autoregressive Conditional Heteroscedasticity (ARCH) models have been widely used in financial time series analysis and particularly in analyzing the risk of holding an asset, evaluating the price of an option, forecasting time varying confidence intervals and obtaining more efficient estimators under the existence of heteroscedasticity.

Let $\{y_t(\theta)\}$ refer to the univariate discrete time real-valued stochastic process to be predicted (e.g. the rate of return of a particular stock or market portfolio from time $t-1$ to t) where θ is a vector of unknown parameters and $E(y_t(\theta)\,|\,I_{t-1}) \equiv E_{t-1}(y_t(\theta)) \equiv \mu_t(\theta)$ denotes the conditional mean given the information

set I_{t-1} (sigma-field) available in time $t-1$. The innovation process for the conditional mean, $\{\varepsilon_t(\theta)\}$, is then given by $\varepsilon_t(\theta) = y_t(\theta) - \mu_t(\theta)$ with corresponding unconditional variance $V(\varepsilon_t(\theta)) = E(\varepsilon_t^2(\theta)) \equiv \sigma^2(\theta)$, zero unconditional mean and $E(\varepsilon_t(\theta)\varepsilon_s(\theta)) = 0$, $\forall t \neq s$. The conditional variance of the process given I_{t-1} is defined by $V(y_t(\theta) \mid I_{t-1}) \equiv V_{t-1}(y_t(\theta)) \equiv E_{t-1}(\varepsilon_t^2(\theta)) \equiv \sigma_t^2(\theta)$. Since investors would know the information set I_{t-1} when they make their investment decisions at time $t-1$, the relevant expected return to the investors and volatility are $\mu_t(\theta)$ and $\sigma_t^2(\theta)$, respectively.

An ARCH process, $\{\varepsilon_t(\theta)\}$, can be presented as:

$$\begin{aligned} \varepsilon_t(\theta) &= z_t \sigma_t(\theta) \\ z_t &\overset{i.i.d.}{\sim} f[E(z_t) = 0, V(z_t) = 1] \\ \sigma_t^2(\theta) &= g(\sigma_{t-1}(\theta), \sigma_{t-2}(\theta), ...; \varepsilon_{t-1}(\theta), \varepsilon_{t-2}(\theta), ...; \upsilon_{t-1}, \upsilon_{t-2}, ...), \end{aligned} \tag{2.1}$$

where $E(z_t) = 0$, $V(z_t) = 1$, $f(.)$ is the density function of z_t, $\sigma_t(\theta)$ is a time-varying, positive and measurable function of the information set at time $t-1$, υ_t is a vector of predetermined variables included in I_t, and $g(.)$ is a linear or nonlinear functional form. By definition, $\varepsilon_t(\theta)$ is serially uncorrelated with mean zero, but with a time varying conditional variance equal to $\sigma_t^2(\theta)$. The conditional variance is a linear or nonlinear function of lagged values of σ_t and ε_t, and predetermined variables $(\upsilon_{t-1}, \upsilon_{t-2}, ...)$ included in I_{t-1}. In the sequel, for notational convenience, no explicit indication of the dependence on the vector of parameters, θ, is given when obvious from the context.

Since very few financial time series have a constant conditional mean of zero, an ARCH model can be presented in a regression form by letting ε_t be the innovation process in a linear regression:

$$\begin{aligned} y_t &= x_t' b + \varepsilon_t \\ \varepsilon_t \mid I_{t-1} &\sim f(0, \sigma_t^2) \\ \sigma_t^2 &= g(\sigma_{t-1}(\theta), \sigma_{t-2}(\theta), ...; \varepsilon_{t-1}(\theta), \varepsilon_{t-2}(\theta), ...; \upsilon_{t-1}, \upsilon_{t-2}, ...), \end{aligned} \tag{2.2}$$

where x_t is a $k \times 1$ vector of endogenous and exogenous explanatory variables included in the information set I_{t-1} and b is a $k \times 1$ vector of unknown parameters.

2.1 ARCH Models

In the literature, one can find a large number of specifications of ARCH models that have been considered for the description of the characteristics of financial markets. A wide range of proposed ARCH processes is covered in surveys such as Bera and Higgins (1993), Bollerslev et al. (1992), Bollerslev et al. (1994), Gouriéroux (1997) and Li et al. (2001).

Engle (1982) introduced the original form of $\sigma_t^2 = g(.)$, in equation (2.1), as a linear function of the past q squared innovations:

$$\sigma_t^2 = a_0 + \sum_{i=1}^{q} \left(a_i \varepsilon_{t-i}^2 \right). \tag{2.3}$$

For the linear ARCH(q) process to be well defined and the conditional variance to be positive, almost surely the parameters must satisfy $\alpha_0 > 0$, $a_i \geq 0$, for $i = 1,..., q$. An equivalent representation of the ARCH(q) process is given by:

$$\sigma_t^2 = a_0 + A(L)\varepsilon_t^2, \tag{2.4}$$

where L denotes the lag operator and $A(L) = \left(a_1 L + a_2 L^2 + ... + a_q L^q \right)$. Defining $v_t = \varepsilon_t^2 - \sigma_t^2$, the model is rewritten as:

$$\varepsilon_t^2 = a_0 + A(L)\varepsilon_t^2 + v_t. \tag{2.5}$$

By its definition, v_t is serially uncorrelated with $E_{t-1}(v_t) = 0$ but neither independently nor identically distributed. The ARCH(q) model is interpreted as an autoregressive process in the squared innovations and is covariance stationary if and only if the roots of $\sum_{i=1}^{q} \left(a_i L^i \right) = 1$ lie outside the unit circle, or, equivalently, the sum of the positive autoregressive parameters is less than one. If the process is covariance stationary, its unconditional variance is equal to $V(\varepsilon_t) \equiv \sigma^2 = a_0 \left(1 - \sum_{i=1}^{q} (a_i) \right)^{-1}$.

Also, by definition, the innovation process is serially uncorrelated but not independently distributed. On the other hand, the standardized innovations are time invariant distributed. Thus, the unconditional distribution for the innovation process will have fatter tails than the distribution for the standardized innovations. For example, consider the kurtosis for the ARCH(1) process with conditional normally distributed

innovations is $E\left(\varepsilon_t^4\right)/E\left(\varepsilon_t^2\right)^2 = 3\left(1-\alpha_1^2\right)/\left(1-3\alpha_1^2\right)$ if $3a_1^2 < 1$, and $E\left(\varepsilon_t^4\right)/E\left(\varepsilon_t^2\right)^2 = \infty$ otherwise, i.e., greater than 3, the kurtosis value of the normal distribution. Generally speaking, an ARCH process always has fatter tails than the normal distribution:

$$E\left(\varepsilon_t^4\right)/E\left(\varepsilon_t^2\right)^2 = E\left(\sigma_t^4 z_t^4\right)/E\left(\sigma_t^2 z_t^2\right)^2 = 3E\left(\sigma_t^4\right)/E\left(\sigma_t^2\right)^2 \geq 3E\left(\sigma_t^2\right)^2/E\left(\sigma_t^2\right)^2 ,$$

where the first equality comes from the independence of σ_t and z_t, and the inequality is implied by Jensen's inequality.

In empirical applications of the ARCH(q) model, a relatively long lag in the conditional variance equation is often called for, and to avoid problems of negative variance parameter estimates a fixed lag structure is typically imposed (see, for example, Engle (1982, 1983), and Engle and Kraft (1983)). To circumvent this problem, Bollerslev (1986) proposed a generalization of the ARCH(q) process to allow for past conditional variances in the current conditional variance equation, the generalized ARCH, or GARCH(p,q), model:

$$\sigma_t^2 = a_0 + \sum_{i=1}^{q}\left(a_i \varepsilon_{t-i}^2\right) + \sum_{j=1}^{p}\left(b_j \sigma_{t-j}^2\right) = a_0 + A(L)\varepsilon_t^2 + B(L)\sigma_t^2 . \qquad (2.6)$$

For $\alpha_0 > 0, a_i \geq 0$, $i = 1,...,q$ and $b_j \geq 0$, $j = 1,...,p$, the conditional variance is well defined. Taylor (1986) independently proposed the GARCH model using a different acronym. Nelson and Cao (1992) showed that the non-negativity constraints on the parameters of the process could be substantially weakened, so they should not be imposed in estimation. Provided that the roots of $B(L) = 1$ lie outside the unit circle and the polynomials $1 - B(L)$ and $A(L)$ have no common roots, the positivity constraint is satisfied if all the coefficients in the infinite power series expansion for $B(L)(1 - B(L))^{-1}$ are non-negative. In the GARCH(1,2) model, for example, the conditions of non-negativity are that $a_0 \geq 0$, $0 \leq b_1 < 1$, $a_1 \geq 0$ and $b_1 a_1 + a_2 \geq 0$. In the GARCH(2,1) model, the necessary conditions require that $a_0 \geq 0$, $b_1 \geq 0$, $a_1 \geq 0$, $b_1 + b_2 < 1$ and $b_1^2 + 4b_2 \geq 0$. Thus, slightly negative values of parameters, for higher order lags, do not result in negative conditional variance. Rearranging the GARCH(p,q) model, it can be presented as an autoregressive moving average process in the squared innovations of orders $\max(p,q)$ and p, $[ARMA(\max(p,q), p)]$, respectively:

$$\varepsilon_t^2 = a_0 + \sum_{i=1}^{q}\left(a_i\varepsilon_{t-i}^2\right) + \sum_{j=1}^{p}\left(b_j\varepsilon_{t-j}^2\right) - \sum_{j=1}^{p}\left(b_j v_{t-j}\right) + v_t . \tag{2.7}$$

The model is second order stationary if the roots of $A(L)+B(L)=1$ lie outside the unit circle, or equivalently if $\sum_{i=1}^{q} a_i + \sum_{j=1}^{p} b_j < 1$. Its unconditional variance is equal to $\sigma^2 = a_0\left(1-\sum_{i=1}^{q} a_i - \sum_{j=1}^{p} b_j\right)^{-1}$.

Very often, in connection with applications, the estimate for $A(L)+B(L)$ turns out to be very close to unity. This provided an empirical motivation, for the development of the so-called integrated GARCH(p,q) or IGARCH(p,q) model by Engle and Bollerslev (1986):

$$\sigma_t^2 = a_0 + A(L)\varepsilon_t^2 + B(L)\sigma_t^2 \text{, for } A(L)+B(L)=1, \tag{2.8}$$

where the polynomial $A(L)+B(L)=1$ has $d>0$ unit roots and $\max(p,q)-d$ roots outside the unit circle.

Moreover, Nelson (1990a) showed that the GARCH(1,1) model is strictly stationary even if $a_1+b_1>1$, as long as $E\left(\log\left(b_1+a_1 z_t^2\right)\right)<0$. Thus, the conditional variance in IGARCH(1,1) with $a_0=0$, collapses to zero almost surely, and in IGARCH(1,1) with $a_0>0$ is strictly stationary. Therefore, a process that is integrated in the mean is not stationary in any sense, while an IGARCH process is strictly stationary but covariance non-stationary.

Consider the IGARCH(1,1) model, $\sigma_t^2 = a_0 + a_1\varepsilon_{t-1}^2 + (1-a_1)\sigma_{t-1}^2$, where $0<a_1<1$. The conditional variance h-steps in the future takes the form:

$$E_t\left(\sigma_{t+h}^2\right) = \sigma_{t+h|t}^2 = \sigma_t^2 + ha_0 , \tag{2.9}$$

which looks very much like a linear random walk with drift a_0. A linear random walk is strictly non-stationary (no stationary distribution and covariance non-stationary) and it has no unconditional first or second moments. In the case of IGARCH(1,1), the conditional variance is strictly stationary even though its stationary distribution generally lacks unconditional moments. In the case where $a_0=0$, equation (2.9) reduces to $\sigma_{t+h|t}^2 \equiv \sigma_t^2$, a bounded martingale as it cannot take negative values. According to the martingale convergence theorem (Dudley (1989)), a bounded martingale must converge, and, in this case, the only value to which it can converge is zero. Thus, the stationary

distributions for σ_t^2 and ε_t have moments, but they are all trivially zero. In the case of $a_0 > 0$, Nelson (1990a) showed that there is a non-degenerate stationary distribution for the conditional variance, but with no finite mean or higher moments. The innovation process ε_t then has a stationary distribution with zero mean, but with tails that are so thick that no second or higher order moments exist. Furthermore, if the variable z_t follows the standard normal distribution, Nelson (1990a) showed that:

$$E\left(\log\left(b_1 + a_1 z_t^2\right)\right) = \log\left(2a_1\right) + \psi\left(1/2\right) + \left(2\pi b_1 a_1^{-1}\right)^{1/2} \Phi\left(0.5;1.5;b_1/2a_1\right) - \left(b_1/a_1\right)_2 F_2\left(1,1;2,1.5;b_1/2a_1\right), \tag{2.10}$$

where $\psi(.)$ denotes the Euler Psi function, with $\psi(1/2) \approx -1.96351$ (Davis (1965)), $\Phi(.;.;.)$ the confluent hypergeometric function (Lebedev (1972)), and ${}_2F_2(.,.;.,.;.)$ the generalized hypergeometric function (Lebedev (1972)). Bougerol and Picard (1992) extended Nelson's work and showed that the general GARCH(p,q) model is strictly stationary and ergodic. Choudhry (1995), by means of the IGARCH(1,1) model, studied the persistence of stock return volatility in European markets during the 1920's and 1930's and argued that the 1929 stock market crash did not reduce stock market volatility. Using monthly stock returns from 1919 to 1936 in markets of Czechoslovakia, France, Italy, Poland and Spain, Choudhry mentioned that in the GARCH(1,1) model the sum of a_1 and b_1 approaches unity, which implies persistence of a forecast of the conditional variance over all finite horizons.

The GARCH(p,q) model successfully captures several characteristics of financial time series, such as thick tailed returns and volatility clustering. On the other hand, its structure imposes important limitations. The variance only depends on the magnitude and not the sign of ε_t, which is somewhat at odds with the empirical behavior of stock market prices where the "leverage effect" may be present. The models that have been considered so far are symmetric in that only the magnitude and not the positivity or negativity of innovations determines σ_t^2. In order to capture the asymmetry manifested by the data, a new class of models, in which good news and bad news have different predictability for future volatility, was introduced.

The most popular method proposed to capture the asymmetric effects is Nelson's (1991) exponential GARCH, or EGARCH, model. He proposed the following form for the evolution of the conditional variance:

$$\log\left(\sigma_t^2\right) = a_0 + \sum_{i=1}^{\infty} \pi_i g\left(\frac{\varepsilon_{t-i}}{\sigma_{t-i}}\right), \quad \pi_1 \equiv 1, \tag{2.11}$$

and accommodated the asymmetric relation between stock returns and volatility changes by making $g\left(\varepsilon_t/\sigma_t\right)$ a linear combination of $\left|\varepsilon_t/\sigma_t\right|$ and ε_t/σ_t :

$$g\left(\varepsilon_t/\sigma_t\right) \equiv \theta\left(\left|\varepsilon_t/\sigma_t\right| - E\left|\varepsilon_t/\sigma_t\right|\right) + \gamma\left(\varepsilon_t/\sigma_t\right), \tag{2.12}$$

where θ and γ are constants. By construction, equation (2.12) is a zero mean i.i.d. sequence (note that $z_t \equiv \varepsilon_t/\sigma_t$). Over the range $0 < z_t < \infty$, $g\left(z_t\right)$ is linear in z_t with slope $\theta + \gamma$ and over the range $-\infty < z_t \le 0$, $g\left(z_t\right)$ is linear with slope $\gamma - \theta$. The first term of (2.12), $\theta\left(\left|z_t\right| - E\left|z_t\right|\right)$, represents the magnitude effect as in the GARCH model, while the second term, $\gamma\left(z_t\right)$, represents the leverage effect. To make this tangible, assume that $\theta > 0$ and $\gamma = 0$. The innovation in $\log\left(\sigma_t^2\right)$ is then positive (negative) when the magnitude of z_t is larger (smaller) than its expected value. Assume now that $\theta = 0$ and $\gamma < 0$. In this case the innovation in $\log\left(\sigma_t^2\right)$ is positive (negative) when innovations are negative (positive). Moreover, the conditional variance is positive regardless of whether the π_i coefficients are positive. Thus, in contrast to GARCH models, no inequality constraints need to be imposed for estimation. Nelson (1991) showed that $\log\left(\sigma_t^2\right)$ and ε_t are strictly stationary as long as $\sum_{i=1}^{\infty} \pi_i^2 < \infty$. A natural parameterization is to model the infinite moving average representation of equation (2.11) as an autoregressive moving average model:

$$\log\left(\sigma_t^2\right) = a_0 + \left(1 + \sum_{i=1}^{q} a_i L^i\right)\left(1 - \sum_{j=1}^{p} b_j L^j\right)^{-1} \left(\theta\left(\left|\varepsilon_{t-1}/\sigma_{t-1}\right| - E\left|\varepsilon_{t-1}/\sigma_{t-1}\right|\right) + \gamma\left(\varepsilon_{t-1}/\sigma_{t-1}\right)\right), \tag{2.13}$$

or equivalently:

$$\log\left(\sigma_t^2\right) = a_0 + \left(1 + A(L)\right)\left(1 - B(L)\right)^{-1} g\left(z_{t-1}\right). \tag{2.13b}$$

Another popular way to model the asymmetry of positive and negative innovations is the use of indicator functions. Glosten et al. (1993) presented the GJR(p,q) model:

$$\sigma_t^2 = a_0 + \sum_{i=1}^{q}\left(a_i \varepsilon_{t-i}^2\right) + \sum_{i=1}^{q}\left(\gamma_i d\left(\varepsilon_{t-i} < 0\right)\varepsilon_{t-i}^2\right) + \sum_{j=1}^{p}\left(b_j \sigma_{t-j}^2\right), \tag{2.14}$$

where γ_i, for $i = 1,...,q$, are parameters that have to be estimated, $d(.)$ denotes the indicator function (i.e. $d(\varepsilon_{t-i} < 0) = 1$ if $\varepsilon_{t-i} < 0$, and $d(\varepsilon_{t-i} < 0) = 0$ otherwise). The GJR model allows good news, $(\varepsilon_{t-i} > 0)$, and bad news, $(\varepsilon_{t-i} < 0)$, to have differential effects on the conditional variance. Therefore, in the case of the GJR(0,1) model, good news has an impact of a_1, while bad news has an impact of $a_1 + \gamma_1$. For $\gamma_1 > 0$, the "leverage effect" exists.

A similar way to model asymmetric effects on the conditional standard deviation was introduced by Zakoian (1990), and developed further in Rabemananjara and Zakoian (1993), by defining the threshold GARCH, or TGARCH(p,q), model:

$$\sigma_t = a_0 + \sum_{i=1}^{q} \left(a_i \varepsilon_{t-i}^{+} \right) - \sum_{i=1}^{q} \left(\gamma_i \varepsilon_{t-i}^{-} \right) + \sum_{j=1}^{p} \left(b_j \sigma_{t-j} \right), \tag{2.15}$$

where $\varepsilon_t^{+} \equiv \varepsilon_t$ if $\varepsilon_t > 0$, $\varepsilon_t^{+} \equiv 0$ otherwise and $\varepsilon_t^{-} \equiv \varepsilon_t - \varepsilon_t^{+}$.

Engle and Ng (1993) recommended the "news impact curve" as a measure of how news is incorporated into volatility estimates by alternative ARCH models. In their recent comparative study of the EGARCH model to the GJR model, Friedmann and Sanddorf-Köhle (2002) proposed a modification of the news impact curve termed the "conditional news impact curve". Engle and Ng argued that the GJR model is better than the EGARCH model because the conditional variance implied by the latter is too high due to its exponential functional form. On the other hand, Friedmann and Sanddorf-Köhle (2002) argued that the EGARCH model does not overstate the predicted volatility.

The number of formulations presented in the financial and econometric literature is vast. In the sequel, the best known variations of ARCH modeling are presented.

Taylor (1986) and Schwert (1989a,b) assumed that the conditional standard deviation is a distributed lag of absolute innovations, and introduced the absolute GARCH, or AGARCH(p,q), model:

$$\sigma_t = a_0 + \sum_{i=1}^{q} a_i \left| \varepsilon_{t-i} \right| + \sum_{j=1}^{p} b_j \sigma_{t-j} \ . \tag{2.16}$$

Geweke (1986), Pantula (1986) and Milhøj (1987) suggested a specification in which the log of the conditional variance depends linearly on past logs of squared innovations. Their model is the multiplicative ARCH, or Log-GARCH(p,q), model defined by

$$\log\left(\sigma_t^2\right) = a_0 + \sum_{i=1}^{q} a_i \log\left(\varepsilon_{t-i}^2\right) + \sum_{j=1}^{p} b_j \log\left(\sigma_{t-j}^2\right). \tag{2.17}$$

Schwert (1990) built the autoregressive standard deviation, or Stdev-ARCH(q), model:

$$\sigma_t^2 = \left(a_0 + \sum_{i=1}^{q} a_i \left| \varepsilon_{t-i} \right| \right)^2 . \tag{2.18}$$

Higgins and Bera (1992) introduced the non-linear ARCH, or NARCH(p,q), model:

$$\sigma_t^{\delta} = a_0 + \sum_{i=1}^{q} a_i \left| \varepsilon_{t-i}^2 \right|^{\delta/2} + \sum_{j=1}^{p} b_j \sigma_{t-j}^{\delta} , \tag{2.19}$$

while Engle and Bollerslev (1986) proposed a simpler non-linear ARCH model:

$$\sigma_t^2 = a_0 + a_1 \left| \varepsilon_{t-1} \right|^{\delta} + b_1 \sigma_{t-1}^2 . \tag{2.20}$$

In order to introduce asymmetric effects, Engle (1990), proposed the asymmetric GARCH, or AGARCH(p,q), model:

$$\sigma_t^2 = a_0 + \sum_{i=1}^{q} \left(a_i \varepsilon_{t-i}^2 + \gamma_i \varepsilon_{t-i} \right) + \sum_{j=1}^{p} b_j \sigma_{t-j}^2 , \tag{2.21}$$

where a negative value of γ_i means that positive returns increase volatility less than negative returns. Moreover, Engle and Ng (1993) presented two more ARCH models that incorporate asymmetry for good and bad news, the non-linear asymmetric GARCH, or NAGARCH(p,q), model:

$$\sigma_t^2 = a_0 + \sum_{i=1}^{q} a_i \left(\varepsilon_{t-i} + \gamma_i \sigma_{t-i} \right)^2 + \sum_{j=1}^{p} b_j \sigma_{t-j}^2 , \tag{2.22}$$

and the VGARCH(p,q) model:

$$\sigma_t^2 = a_0 + \sum_{i=1}^{q} a_i \left(\varepsilon_{t-i} / \sigma_{t-i} + \gamma_i \right)^2 + \sum_{j=1}^{p} b_j \sigma_{t-j}^2 . \tag{2.23}$$

Ding et al. (1993) introduced the asymmetric power ARCH, or APARCH(p,q), model, which includes seven ARCH models as special cases (ARCH, GARCH, AGARCH, GJR, TARCH, NARCH and logARCH):

$$\sigma_t^{\delta} = a_0 + \sum_{i=1}^{q} a_i \left(\left| \varepsilon_{t-1} \right| - \gamma_i \varepsilon_{t-i} \right)^{\delta} + \sum_{j=1}^{p} b_j \sigma_{t-j}^{\delta} , \tag{2.24}$$

where $a_0 > 0$, $\delta \geq 0$, $b_j \geq 0$, $j = 1,..., p$, $a_i \geq 0$ and $-1 < \gamma_i < 1$, $i = 1,..., q$. The model imposes a Box and Cox (1964) power transformation of the conditional standard deviation process and the asymmetric absolute innovations. The functional form for the conditional standard deviation is familiar to economists as the constant elasticity of substitution (CES) production function. Ling and McAleer (2001) provided sufficient

conditions for the stationarity and ergodicity of the APARCH(p,q), model. Brooks et al. (2000) applied the APARCH(1,1) model for 10 series of national stock market index returns. The optimal power transformation was found to be remarkably similar across countries.

Sentana (1995) introduced the quadratic GARCH, or GQARCH(p,q), model of the form:

$$\sigma_t^2 = a_0 + \sum_{i=1}^{q} a_i \varepsilon_{t-i}^2 + \sum_{i=1}^{q} \gamma_i \varepsilon_{t-i} + 2\sum_{i=1}^{q}\sum_{j=i+1}^{q} a_{ij}\varepsilon_{t-i}\varepsilon_{t-j} + \sum_{j=1}^{p} b_j \sigma_{t-j}^2 \,. \tag{2.25}$$

Setting $\gamma_i = 0$, for $i = 1,...,q$, leads to the Augmented ARCH model of Bera and Lee (1990). It does encompass all the ARCH models of quadratic variance functions, but it does not include models in which the variance is quadratic in the absolute value of innovations, as the APARCH model.

Hentschel (1995) gave a complete parametric family of ARCH models. This family nests the most popular symmetric and asymmetric ARCH models, thereby highlighting the relation between the models and their treatment of asymmetry. Hentschel presents the variance equation as:

$$\frac{\sigma_t^{\lambda} - 1}{\lambda} = \omega + a\sigma_{t-1}^{\lambda} f^{v}(\varepsilon_t) + \beta \frac{\sigma_{t-1}^{\lambda} - 1}{\lambda} \,, \tag{2.26}$$

where $f(.)$ denotes the absolute value function of innovations,

$$f(\varepsilon_t) = |\varepsilon_t - \beta| - \zeta(\varepsilon_t - \beta) \,. \tag{2.27}$$

In general, this is a law of the Box-Cox transformation of the conditional standard deviation (as in the case of the APARCH model), and the parameter λ determines the shape of the transformation. For $\lambda > 1$, the transformation of σ_t is convex, while for $\lambda < 1$, it is concave. The parameter v serves to transform the absolute value function. For different restrictions on the parameters in equations (2.26) and (2.27), almost all the popular symmetric and asymmetric ARCH models are obtained. For example, for $\lambda = 0$, $v = 1$, $\beta = 1$ and free ζ, we obtain Nelson's exponential GARCH model. However, some models, as Sentana's quadratic model, are excluded.

Gouriéroux and Monfort (1992) proposed the qualitative threshold GARCH, or GQTARCH(p,q), model with the following specification:

$$\sigma_t^2 = \omega + \sum_{i=1}^{q}\sum_{j=1}^{J} a_{ij} I_j(\varepsilon_{t-i}) + \sum_{i=1}^{p} b_j \sigma_{t-j}^2 \,. \tag{2.28}$$

Assuming constant conditional variance over various observation intervals, Gouriéroux and Monfort (1992) divided the space of ε_t into J intervals and let $I_j(\varepsilon_t)$ be 1 if ε_t is in the j^{th} interval.

Another important class of models, proposed independently by Cai (1994) and Hamilton and Susmel (1994), is the class of regime switching ARCH models, a natural extension of regime-switching models for the conditional mean, introduced by Hamilton (1989). These models allow the parameters of the ARCH process to come from one of several different regimes, with transitions between regimes governed by an unobserved Markov chain. Let $\tilde{\varepsilon}_t$ be the innovation process and let s_t denote an unobserved random variable that can take on the values $1,2,...,K$. Suppose that s_t can be described by a Markov chain, $P(s_t = j \mid s_{t-1} = i, s_{t-2} = k,..., \tilde{\varepsilon}_{t-1}, \tilde{\varepsilon}_{t-2},...) = p_{ij}$, for $i, j = 1,2,...,K$. The idea is to model the innovation process, $\tilde{\varepsilon}_t$, as $\tilde{\varepsilon}_t \equiv \sqrt{g_{s_t}}\varepsilon_t$, where ε_t is assumed to follow an ARCH process. So, the underlying ARCH variable, ε_t, is multiplied by the constant $\sqrt{g_1}$ when the process is in the regime presented by $s_t = 1$, is multiplied by $\sqrt{g_2}$ when $s_t = 2$, and so on. The factor for the first stage, g_1, is normalized at unity with $g_j \geq 1$ for $j = 2,3,...,K$. The idea is, thus, to model changes in regime as changes in the scale of the process. Dueker (1997) and Hansen (1994) extended the approach to GARCH models.

Fornari and Mele (1995) introduced the volatility-switching ARCH model, or VSARCH(p,q), model:

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} a_i \varepsilon_{t-i}^2 + \gamma S_{t-1} \frac{\varepsilon_{t-1}^2}{\sigma_{t-1}^2} + \sum_{j=1}^{p} b_j \sigma_{t-j}^2 , \qquad (2.29)$$

where S_t is an indicator factor that equals one if $\varepsilon_t > 0$, minus one if $\varepsilon_t < 0$, and $\varepsilon_t^2 / \sigma_t^2$ measures the difference between the forecast of the volatility at time t on the basis of the information set dated at $t-1$, σ_t^2, and the realized value ε_t^2. As Fornari and Mele (1995) mentioned, the volatility-switching model is able to capture a phenomenon that has not been modeled before. It implies that asymmetries can become inverted, with positive innovations inducing more volatility than negative innovations of the same size when the observed value of the conditional variance is lower than expected. Fornari and

Mele (1996) built a mixture of the GJR and the VSARCH models, named it asymmetric volatility-switching ARCH, or AVSARCH(p,q), model and estimated it for $p = q = 1$:

$$\sigma_t^2 = a_0 + a_1\varepsilon_{t-1}^2 + b_1\sigma_{t-1}^2 + \gamma S_{t-1}\varepsilon_{t-1}^2 + \delta\left(\left(\varepsilon_{t-1}^2/\sigma_{t-1}^2\right) - k\right)S_{t-1}. \quad (2.30)$$

The first four terms are the GJR(1,1) model, except that S_t is a dummy that equals one or minus one instead of zero or one, respectively. The last term captures the reversal of asymmetry observed when $\varepsilon_{t-1}^2/\sigma_{t-1}^2$ reaches k, the threshold value. Note that the AVSARCH model is able to generate kurtosis higher than the GARCH or GJR models.

Hagerud (1996), inspired by the Smooth Transition Autoregressive (STAR) model of Luukkonen et al. (1988), proposed the smooth transition ARCH model. In the STAR model, the conditional mean is a non-linear function of lagged realizations of the series introduced via a transition function. The smooth transition GARCH(p,q) model has the form:

$$\sigma_t^2 = a_0 + \sum_{i=1}^{q}\left(a_i + \gamma_i F\left(\varepsilon_{t-i}\right)\right)\varepsilon_{t-i}^2 + \sum_{j=1}^{p} b_j\sigma_{t-j}^2, \quad (2.31)$$

where $F(.)$ is either the logistic or the exponential transition function, the two most commonly used transition functions for STAR models (for details see Teräsvirta (1994)). The logistic function considered is

$$F\left(\varepsilon_{t-i}\right) = \left(1 + \exp\left(-\theta\varepsilon_{t-i}\right)\right)^{-1} - 0.5, \text{ for } \theta > 0, \quad (2.32)$$

and the exponential function is

$$F\left(\varepsilon_{t-i}\right) = 1 - \exp\left(-\theta\varepsilon_{t-i}^2\right), \text{ for } \theta > 0. \quad (2.33)$$

The two resulting models termed logistic and exponential smooth transition GARCH, or LST-GARCH(p,q) and EST-GARCH(p,q), models, respectively. The smooth transition models allow for the possibility of intermediate positions between different regimes. For $-\infty < \varepsilon_t < \infty$, the logistic transition function takes values in $-0.5 \le F(.) \le 0.5$ and generates data where the dynamics of the conditional variance differ depending on the sign of innovations. On the other hand, the exponential function generates a return process for which the dynamics of the conditional variance depend on the magnitude of the innovations, as for $|\varepsilon_t| \to \infty$ the transition function will be equal to unity, and when $\varepsilon_t = 0$ the transition function is equal to zero. Thus, contrary to the regime switching models, the transition between states is smooth as the conditional variance is a continuous function of innovations. A model similar to the LST-GARCH model was

independently proposed by González-Rivera (1996). Recently, Nam et al. (2002) provided an application of a smooth transition ARCH model with a logistic function in the following form

$$\sigma_t^2 = a_0 + a_1\varepsilon_{t-1}^2 + a_2\sigma_{t-1}^2 + \left(b_0 + b_1\varepsilon_{t-1}^2 + b_2\sigma_{t-1}^2\right)F\left(\varepsilon_{t-1}\right)$$
$$F\left(\varepsilon_{t-1}\right) = \left(1+\exp\left(-\theta\varepsilon_{t-1}\right)\right)^{-1},$$

which they termed asymmetric nonlinear smooth transition GARCH, or ANST-GARCH model. Nam et al. explored the asymmetric reverting property of short-horizon expected returns and have found that the asymmetric return reversals can be exploited for the contrarian profitability[1]. Note that when $b_0 = b_2 = 0$ the ANST-GARCH model reduces to González-Rivera's specification. Lubrano (1998) suggested an improvement over these transition functions, introducing an extra parameter, the threshold c, which determines at which magnitude of past innovations the change of regime occurs. The generalized logistic transition function is given by:

$$F\left(\varepsilon_{t-i}\right) = \frac{1-\exp\left(-\theta\varepsilon_{t-i}^2\right)}{1+\exp\left(-\theta\left(\varepsilon_{t-i}^2 - c^2\right)\right)}. \tag{2.34}$$

The exponential transition function can also be generalized in the form:

$$F\left(\varepsilon_{t-i}\right) = 1-\exp\left(-\theta\left(\varepsilon_{t-i} - c\right)^2\right). \tag{2.35}$$

Engle and Lee (1993) proposed the component GARCH model in order to investigate the long-run and the short-run movement of volatility. The GARCH(1,1) model can be written as:

$$\sigma_t^2 = \sigma^2 + a_1\left(\varepsilon_{t-1}^2 - \sigma^2\right) + b_1\left(\sigma_{t-1}^2 - \sigma^2\right), \tag{2.36}$$

for $\sigma^2 = a_0\left(1-a_1-b_1\right)^{-1}$ denoting the unconditional variance. The conditional variance in the GARCH(1,1) model shows mean reversion to the unconditional variance, which is constant for all time. By contrast, the component GARCH, or CGARCH(1,1), model allows mean reversion to a time varying level q_t. The CGARCH(1,1) model is defined as:

[1] Contrarian investment strategies are contrary to the general market direction. Interpretation of the contrarian profitability is in a debate between the two competing hypotheses: the time varying rational expectation hypothesis and the stock market overreaction hypothesis. For details see Chan (1988), Chopra et al. (1992), Conrad and Kaul (1993), DeBondt and Thaler (1985, 1987,1989), Lo and MacKinlay (1990b), Veronesi (1999), Zarowin (1990).

$$\begin{aligned}\sigma_t^2 &= q_t + a_1\left(\varepsilon_{t-1}^2 - q_{t-1}\right) + b_1\left(\sigma_{t-1}^2 - q_{t-1}\right)\\ q_t &= a_0 + pq_{t-1} + \phi\left(\varepsilon_{t-1}^2 - \sigma_{t-1}^2\right).\end{aligned} \tag{2.37}$$

The difference between the conditional variance and its trend, $\sigma_t^2 - q_t$, is the transitory or short-run component of the conditional variance, while q_t is the time varying long-run volatility. Combining the transitory and permanent equations the model reduces to:

$$\begin{aligned}\sigma_t^2 &= (1-a_1-b_1)(1-p)a_0 + (a_1+\phi)\varepsilon_{t-1}^2 - (a_1 p + (a_1+b_1)\phi)\varepsilon_{t-2}^2\\ &\quad + (b_1-\phi)\sigma_{t-1}^2 - (b_1 p - (a_1+b_1)\phi)\sigma_{t-2}^2,\end{aligned} \tag{2.38}$$

which shows that the CGARCH(1,1) is a restricted GARCH(2,2) model. Moreover, because of the existence of the "leverage effect", Engle and Lee (1993) combine the component model with the GJR model to allow shocks to affect the volatility components asymmetrically. The asymmetric component GARCH, or the ACGARCH(1,1), model becomes:

$$\begin{aligned}\sigma_t^2 &= q_t + a_1\left(\varepsilon_{t-1}^2 - q_{t-1}\right) + \gamma_1\left(d(\varepsilon_{t-1}<0)\varepsilon_{t-1}^2 - 0.5q_{t-1}\right) + b_1\left(\sigma_{t-1}^2 - q_{t-1}\right)\\ q_t &= a_0 + pq_{t-1} + \phi\left(\varepsilon_{t-1}^2 - \sigma_{t-1}^2\right) + \gamma_2\left(d(\varepsilon_{t-1}<0)\varepsilon_{t-1}^2 - 0.5\sigma_{t-1}^2\right),\end{aligned} \tag{2.39}$$

where $d(.)$ denotes the indicator function (i.e. $d(\varepsilon_{t-i}<0)=1$ if $\varepsilon_{t-i}<0$, and $d(\varepsilon_{t-i}<0)=0$ otherwise).

Baillie et al. (1996), motivated by the Fractionally Integrated Autoregressive Moving Average, or ARFIMA, model, presented the Fractionally Integrated Generalized Autoregressive Conditional Heteroscedasticity, or FIGARCH, model. The ARFIMA(k,d,l) model for the discrete time real-valued process $\{y_t\}$, initially developed in Granger (1980) and Granger and Joyeux (1980), is defined as:

$$A(L)(1-L)^d\, y_t = B(L)\varepsilon_t, \tag{2.40}$$

where $A(L)$ and $B(L)$ denote the lag operators of order k and l respectively, and $\{\varepsilon_t\}$ is a mean-zero serially uncorrelated process. The fractional differencing operator, $(1-L)^d$, is usually interpreted in its binomial expansion given by:

$$(1-L)^d = \sum_{j=0}^{\infty}\pi_j L^j \text{ , for } \pi_j = \frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)} = \prod_{k=0}^{j}\frac{k-1-d}{k}, \tag{2.41}$$

where $\Gamma(.)$ denotes the gamma function.

The stationary ARMA process, equation (2.40) for $d=0$, is a short memory process, the autocorrelations of which are geometrically bounded:

$$\left|Cor\left(y_t, y_{t+m}\right)\right| \le cr^m,$$

for $m = 1,2,...$, where $c > 0$ and $0 < r < 1$. As $m \to \infty$ the dependence, or memory, between y_t and y_{t+m} decreases rapidly. However, some observed time series appeared to exhibit a substantially larger degree of persistence than allowed for by stationary ARMA processes. For example, Ding et al. (1993) found that the absolute values or powers, particularly squares, of returns on S&P500 index tend to have very slowly decaying autocorrelations. Similar evidence of this feature for other types of financial series is contained in Dacarogna et al. (1993), Mills (1996) and Taylor (1986). Such time series have autocorrelations that seem to satisfy the condition:

$$Cor\left(y_t, y_{t+m}\right) \approx cm^{2d-1},$$

as $m \to \infty$, where $c \neq 0$ and $d < 0.5$. Such processes are said to have long memory because the autocorrelations display substantial persistence.

The concept of long memory and fractional Brownian motion was originally developed by Hurst (1951) and extended by Mandelbrot (1963, 1982) and Mandelbrot and Van Ness (1968). However, the ideas became essentially applicable by Granger (1980,1981), Granger and Joyeux (1980) and Hosking (1981). Hurst was a hydrologist who worked on the Nile river dam project. He had studied an 847-years record of the Nile's overflows and observed that larger than average overflows were more likely to be followed by more large overflows. Suddenly, the water flow would change to a lower than average overflow which would be followed by lower than average overflows. Such a process could be examined neither with standard statistical correlation analysis nor by assuming that the water inflow is a random process, so it could be analyzed as a Brownian motion. Einstein (1905) worked on Brownian motion and found that the distance a random particle covers increases with the square root of time used to measure it, or:

$$d = t^{1/2}, \tag{2.42}$$

where d is the distance covered and t is the time index. But this applies only to time series that are in Brownian motion, i.e. mean-zero and unity variance independent processes. Hurst generalized (2.42) to account for processes other than Brownian motion in the form:

$$d/s = ct^H. \tag{2.43}$$

For any process $\{y_t\}_{t=1}^{T}$ (e.g. asset returns) with mean $\bar{y}_T = T^{-1}\sum_{t=1}^{T} y_t$, d is given by

$$d = \underset{1\le k\le T}{Max}\sum_{t=1}^{k}(y_t - \bar{y}_T) - \underset{1\le k\le T}{Min}\sum_{t=1}^{k}(y_t - \bar{y}_T), \tag{2.44}$$

where s is the standard deviation of $\{y_t\}_{t=1}^{T}$ and c is a constant. The ratio d/s is called *rescaled range* and H is the Hurst exponent. If $\{y_t\}$ is a sequence of independently and identically distributed random variables, then $H = 0.5$. Hurst's investigations for the Nile lead to $H = 0.9$. Thus, the rescaled range was increasing at a faster rate than the square root of time.

The IGARCH(p,q) model in equation (2.8) could be rewritten as:

$$\Phi(L)(1-L)\varepsilon_t^2 = a_0 + (1 - B(L))v_t, \tag{2.45}$$

where $\Phi(L) \equiv (1 - A(L) - B(L))(1-L)^{-1}$ is of order $[\max(p,q) - 1]$. The FIGARCH model is simply obtained by replacing the first difference operator in equation (2.45) with the fractional differencing operator. Rearranging terms in equation (2.45) the FIGARCH(p,d,q) model is given as:

$$\sigma_t^2 = a_0 + \left(1 - B(L) - \Phi(L)(1-L)^d\right)\varepsilon_t^2 + B(L)\sigma_t^2, \tag{2.46}$$

which is strictly stationary and ergodic for $0 \le d \le 1$. In contrast to the GARCH and IGARCH models where shocks to the conditional variance either dissipate exponentially or persist indefinitely, for the FIGARCH model the response of the conditional variance to past shocks decays at a slow hyperbolic rate. The sample autocorrelations of the daily absolute returns, or $|y_t|$, as investigated by Ding et al. (1993) and Bollerslev and Mikkelsen (1996) among others, exceed the 95% confidence intervals for no serial dependence for more than 1000 lags. Moreover, the sample autocorrelations for the first difference of absolute returns, $(1-L)|y_t|$, still show statistically significant long-term dependence. On the contrary, the fractional difference of absolute returns, $(1-L)^{0.5}|y_t|$, shows much less long-term dependence. Bollerslev and Mikkelsen (1996) provided evidence that illustrates the importance of using fractional integrated conditional variance models in the context of pricing options with maturity time of one year or longer. Note that the practical importance of the fractional integrated variance models stems from the added flexibility when modeling long run volatility characteristics.

As Mills (1999) stated, the implication of IGARCH models that shocks to the conditional variance persist indefinitely does not reconcile with the persistence observed after large shocks, such as the crash of October 1987, and with the perceived behavior of agents who do not appear to frequently and radically alter the composition of their portfolios. So the widespread observation of the IGARCH behavior may be an artifact of a long memory FIGARCH data generating process. Baillie et al. (1996) provided a simulation experiment that provides considerable support of this line of argument. Beine et al. (2002) applied the FIGARCH(1,d,1) model in order to investigate the effects of official interventions on the volatility of exchange rates. One of their interesting remarks is that measuring the volatility of exchange rates through the FIGARCH model instead of a traditional ARCH model leads to different results. The GARCH and IGARCH models tend to underestimate the effect of the central bank interventions on the volatility of exchange rates. Vilasuso (2002) fitted conditional volatility models to daily spot exchange rates and found that the FIGARCH(1,d,1) model generates superior volatility forecasts compared to those generated by a GARCH(1,1) or IGARCH(1,1) model.

Bollerslev and Mikkelsen (1996) extended the idea of fractional integration to the exponential GARCH model, whereas Tse (1998) built the fractional integration form of the APARCH model. Factorizing the autoregressive polynomial $(1-B(L))=\Phi(L)(1-L)^d$, where all the roots of $\Phi(z)=0$ lie outside the unit circle, the fractionally integrated exponential GARCH, or FIEGARCH(p,d,q), model is defined as:

$$\log\left(\sigma_t^2\right)=a_0+\Phi(L)^{-1}(1-L)^{-d}\left(1+A(L)\right)g\left(z_{t-1}\right). \tag{2.47}$$

The fractionally integrated asymmetric power ARCH, or FIAPARCH(p,d,q), model has the following form:

$$\sigma_t^{\delta}=a_0+\left(1-\left(1-B(L)\right)^{-1}\Phi(L)(1-L)^{-d}\right)\left(|\varepsilon_t|-\gamma\varepsilon_t\right)^{\delta}. \tag{2.48}$$

Finally, Hwang (2001) presented the asymmetric fractionally integrated family GARCH(1,d,1), or ASYMM FIFGARCH(1,d,1), model, which is defined as:

$$\begin{aligned}\sigma_t^{\lambda}&=\frac{k}{1-\delta}+\left(1-\frac{(1-\varphi L)(1-L)^d}{1-\delta L}\right)f^{v}\left(\varepsilon_t\right)\sigma_t^{\lambda}\\ f\left(\varepsilon_t\right)&=\left|\frac{\varepsilon_t}{\sigma_t}-b\right|-c\left(\frac{\varepsilon_t}{\sigma_t}-b\right),\end{aligned} \tag{2.49}$$

for $|c|\le 1$. Hwang points out that, for different parameter values in (2.49), the following fractionally integrated ARCH models are obtained: FIEGARCH, for $\lambda=0$, $v=1$,

FITGARCH for $\lambda = 1$, $v = 1$, FIGARCH for $\lambda = 2$, $v = 2$, and FINGARCH, for $\lambda = v$ but otherwise unrestricted.

However, Ruiz and Pérez (2002) noted that Hwang's model is poorly specified and does not nest the FIEGARCH model. Thus, they suggested an alternative specification, which is a direct generalization of Hentschel's model in (2.26):

$$\begin{aligned} &(1-\varphi L)(1-L)^{d}\,\frac{\sigma_t^{\lambda}-1}{\lambda} = \omega' + a(1+\psi L)\sigma_{t-1}^{\lambda}\left(f^{v}(z_{t-1})-1\right) \\ &f\left(\frac{\varepsilon_t}{\sigma_t}\right) = \left|\frac{\varepsilon_t}{\sigma_t} - b\right| - c\left(\frac{\varepsilon_t}{\sigma_t} - b\right) \end{aligned}, \qquad (2.50)$$

Imposing appropriate restrictions on the parameters of (2.50), a number of models are obtained as special cases (e.g. the FIGARCH model in (2.46), the FIEGARCH model in (2.47), Hentschel's model in (2.26)).

Nowicka-Zagrajek and Weron (2001) replaced the constant term in the GARCH(p,q) model with a linear function of i.i.d. stable random variables and defined the randomized GARCH, or R-GARCH(r,p,q), model:

$$\sigma_t^2 = \sum_{i^*=1}^{r}\left(c_{i^*}\eta_{t-i^*}\right) + \sum_{i=1}^{q}\left(a_i\varepsilon_{t-i}^2\right) + \sum_{j=1}^{p}\left(b_j\sigma_{t-j}^2\right), \qquad (2.51)$$

where $c_{i^*} \geq 0$, $i^* = 1,...,r$, $a_i \geq 0$, $i = 1,...,q$, $b_j \geq 0$, $j = 1,...,p$, the innovations η_t are positive i.i.d. stable random variables expressed by the characteristic function in (4.16), and $\{\eta_t\}$ and $\{z_t\}$ are independent.

Müller et al. (1997), based on the hypothesis that participants in a heterogeneous market make volatilities of different time resolutions behave differently, proposed the heterogeneous interval GARCH, or H-GARCH(p,n), model that takes into account the squared price changes over time intervals of different sizes:

$$\sigma_t^2 = a_0 + \sum_{i=1}^{n}\sum_{k=1}^{i} a_{ik}\left(\sum_{i^*=k}^{i}\varepsilon_{t-i^*}^2\right)^2 + \sum_{j=1}^{p}\left(b_j\sigma_{t-j}^2\right), \qquad (2.52)$$

where $a_0 > 0$, $a_{ik} \geq 0$, for $i = 1,...,n$, $k = 1,...,i$, $b_j \geq 0$, $j = 1,...,p$.

Many financial markets impose restrictions on the maximum allowable daily change in price. As pointed out by Wei and Chiang (1997), the common practice of ignoring the problem by treating the observed censored observations as if they were actually the equilibrium prices, or dropping the limited prices from the studied sample, leads to the underestimation of conditional volatility. Morgan and Trevor (1997) proposed

the Rational Expectation (RE) algorithm (which can be interpreted as an EM algorithm (Dempster et al. (1977)) for censored observations in the presence of heteroscedasticity, which replaces the unobservable components of the likelihood function of the ARCH model by their rational expectations. As an alternative to the RE algorithm, Wei (2002), based on Kodres's (1993) study, proposed a censored-GARCH model and developed a Bayesian estimation procedure for the proposed model. Moreover, on the basis of Kodres's (1988) research, Lee (1999), Wei (1999) and Calzolari and Fiorentini (1998) developed the class of Tobit-GARCH models.

Brooks et al. (2001) reviewed the most known software packages for estimation of ARCH models, and concluded that the estimation results differ considerably from one another. Table 2, in the Appendix, contains the ARCH models that have been presented in this section.

3. The Relationship Between Conditional Variance and Conditional Mean

3.1 The ARCH in Mean Model

Financial theory suggests that an asset with a higher expected risk would pay a higher return on average. Let y_t denote the rate of return of a particular stock or market portfolio from time t to $t-1$ and rf_t be the return on a riskless asset (i.e. treasury bills). Then, the excess return (asset return minus the return on a riskless asset) can be decomposed into a component anticipated by investors at time $t-1$, μ_t, and a component that was unanticipated, ε_t:

$$y_t - rf_t = \mu_t + \varepsilon_t .$$

The relationship between investors' expected return and risk was presented in an ARCH framework, by Engle et al. (1987). They introduced the ARCH in mean, or ARCH-M, model where the conditional mean is an explicit function of the conditional variance of the process in framework (2.1). The estimated coefficient on the expected risk is a measure of the risk-return tradeoff. Thus, the ARCH regression model, in framework (2.2), can be presented as:

$$y_t = x_t' b + \phi\left(\sigma_t^2\right) + \varepsilon_t$$
$$\varepsilon_t \mid I_{t-1} \sim f\left[0, \sigma_t^2\right]$$
$$\sigma_t^2 = g\left(\sigma_{t-1}, \sigma_{t-2}, ...; \varepsilon_{t-1}, \varepsilon_{t-2}, ...; \upsilon_{t-1}, \upsilon_{t-2}, ...\right).$$

where $\phi\left(\sigma_t^2\right)$ represents the risk premium, i.e., the increase in the expected rate of return due to an increase in the variance of the return. Although earlier studies concentrated on detecting a constant risk premium, the ARCH in mean model provided a new approach by which a time varying risk premium can be estimated. The most commonly used specifications of the ARCH-M model are in the form:

$\phi\left(\sigma_t^2\right) = c_0 + c_1 \sigma_t^2$, (Nelson (1991), Bollerslev et al. (1994)),

$\phi\left(\sigma_t^2\right) = c_0 + c_1 \sigma_t$, (Domowitz and Hakkio (1985), Bollerslev et al. (1988)),

$\phi\left(\sigma_t^2\right) = c_0 + c_1 \log\left(\sigma_t^2\right)$, (Engle et al. (1987)).

A positive as well as a negative risk return tradeoff could be consistent with the financial theory. A positive relationship is expected if we assume a rational risk averse investor who requires a larger risk premium during the times when the payoff of the security is riskier. On the other hand, a negative relationship is expected under the assumption that during relatively riskier periods the investors may want to save more. In applied research work, there is evidence for both positive and negative relationship. French et al. (1987) found positive risk return tradeoff for the excess returns on the S&P500 composite portfolio although not statistically significant in all the examined periods. Nelson (1991) found a negative but insignificant relationship for the excess returns on the Center for Research in Security Prices (CRSP) value weighted market index. Bollerslev et al. (1994) found a positive, not always statistically significant, relationship for the returns on Dow Jones and S&P500 indices. Interesting studies employing the ARCH-M model were conducted by Devaney (2001) and Elyasiani and Mansur (1998). The former examined the tradeoff between conditional variance and excess returns for stocks of the commercial bank sector, while the latter investigated the time varying risk premium for real estate investment trusts.

3.2 Volatility and Serial Correlation

LeBaron (1992) found a strong inverse relation between volatility and serial correlation for S&P500, CRSP value weighted market index, Dow Jones and IBM returns. He introduced the exponential autoregressive GARCH, or EXP-GARCH(p,q), model in which the conditional mean is a non-linear function of the conditional variance. Based on LeBaron (1992), the ARCH regression model, in framework (2.2), can be presented as:

$$\begin{aligned} y_t &= x_t' b + \left(c_1 + c_2 \exp\left(-\sigma_t^2 / c_3\right)\right) y_{t-1} + \varepsilon_t \\ \varepsilon_t \mid I_{t-1} &\sim f\left[0, \sigma_t^2\right] \\ \sigma_t^2 &= g\left(\sigma_{t-1}, \sigma_{t-2}, ...; \varepsilon_{t-1}, \varepsilon_{t-2}, ...; \upsilon_{t-1}, \upsilon_{t-2}, ...\right). \end{aligned} \tag{3.1}$$

The model is a mixture of the GARCH model and the exponential AR model of Ozaki (1980). For the data set LeBaron used, c_2 is significantly negative and remarkably robust to the choice of sample period, market index, measurement interval and volatility measure. As LeBaron stated, it is difficult to estimate c_3 in conjunction with c_2 when using a gradient type of algorithm. So, c_3 is set to the sample variance of the series.

Generally, the first order autocorrelations are larger for periods of lower volatility and smaller during periods of higher volatility. The accumulation of news[2] and the non-synchronous trading[3] were mentioned as the possible reasons. The stocks do not trade close to the end of the day and information arriving during this period is reflected on the next day's trading, inducing serial correlation. As new information reaches market very slowly, traders optimal action is to do nothing until enough information is accumulated. Because of the non-trading, the trading volume, which is strongly positive related with volatility, lowers. Thus, we have a market with low trade volume and high correlation.

Kim (1989), Sentana and Wadhwani (1991) and Oedegaard (1991) have also investigated the relationship between autocorrelation and volatility and found an inverse relation between volatility and autocorrelation. Moreover, Oedegaard (1991) found that the evidence of autocorrelation, for the S&P500 daily index, decreased over time, possibly because of the introduction of financial derivatives (options and futures) on the index.

4. Estimation

4.1 Maximum Likelihood Estimation

In ARCH models, the most commonly used method in estimating the vector of unknown parameters, θ, is the method of maximum likelihood (MLE). Under the assumption of independently and identically distributed standardized innovations, $z_t(\theta) \equiv \varepsilon_t(\theta)/\sigma_t(\theta)$, in framework (2.2), let us denote their density function as $f(z_t; w)$, where $w \in W \subseteq R^{\breve{w}}$ is the vector of the parameters of f to be estimated. So, for $\psi' = (\theta', w')$ denoting the whole set of the $\breve{\psi} = \breve{\theta} + \breve{w}$ parameters that have to be estimated for the conditional mean, variance and density function, the log-likelihood function for $\{y_t(\theta)\}$ is:

$$l_t(y_t; \psi) = \log\left(f\left(z_t(\theta); w\right)\right) - \frac{1}{2}\log\left(\sigma_t^2(\theta)\right). \qquad (4.1)$$

The full sample log-likelihood function for a sample of T observations is simply:

---

[2] See section 1.2.
[3] See section 1.3.

$$L_T\left(\{y_t\};\psi\right)=\sum_{t=1}^{T} l_t\left(y_t;\psi\right). \tag{4.2}$$

If the conditional density, the mean and the variance functions are differentiable for each possible $\psi \in \Theta \times W \equiv \Psi \subseteq R^{\bar{\psi}}$, the MLE estimator $\hat{\psi}$ for the true parameter vector ψ_0 is found by maximizing equation (4.2), or equivalently by solving the equation

$$\sum_{t=1}^{T} \frac{\partial l_t\left(y_t;\psi\right)}{\partial \psi}=0. \tag{4.3}$$

If the density function does not require the estimation of any parameter, as in the case of the normal distribution that is uniquely determined by its first two moments, then $\bar{w}=0$. In such cases, equation (4.3) becomes:

$$\sum_{t=1}^{T}\left(f\left(z_t(\theta)\right)^{-1}\frac{\partial f\left(z_t(\theta)\right)}{\partial \theta}\left(-\frac{\partial \mu_t(\theta)}{\partial \theta}\sigma_t^2(\theta)^{-1}-0.5\sigma_t^2(\theta)^{-3/2}\varepsilon_t(\theta)\frac{\partial \sigma_t^2(\theta)}{\partial \theta}\right)\right)=0. \tag{4.4}$$

Let us, for example, estimate the parameters of framework (2.2) for normal distributed innovations and the GARCH(p,q) functional form for the conditional variance as given in equation (2.6). The density function of the standard normal distribution is:

$$f\left(z_t\right)=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{z_t^2}{2}\right). \tag{4.5}$$

For convenience equation (2.6) is written as $\sigma_t^2=\omega' s_t$, where $\omega'=\left(a_0,a_1,...,a_q,b_1,...,b_p\right)$ and $s_t=\left(1,\varepsilon_{t-1}^2,...,\varepsilon_{t-q}^2,\sigma_{t-1}^2,...,\sigma_{t-p}^2\right)$. The vector of parameters that have to be estimated is $\psi'=\theta'=\left(b',\omega'\right)$. For normally distributed standardized innovations, z_t, the log-likelihood function in equation (4.1), is:

$$l_t\left(y_t;\psi\right)=\log\left(\frac{1}{\sqrt{2\pi}}\right)-\frac{\left(y_t-x_t'b\right)^2}{2\sigma_t^2}-\frac{1}{2}\log\left(\sigma_t^2\right),$$

and the full sample log-likelihood function in equation (4.2), becomes:

$$L_T\left(\{y_t\};\theta\right)=-\frac{1}{2}\left(T\log(2\pi)+\sum_{t=1}^{T}\frac{\left(y_t-x_t'b\right)^2}{\sigma_t^2}+\sum_{t=1}^{T}\log\left(\sigma_t^2\right)\right).$$

The first and the second derivatives of the log-likelihood for the t^{th} observation with respect to the variance parameter vector are:

$$\frac{\partial l_t(y_t;b,\omega)}{\partial \omega} = \frac{1}{2\sigma_t^2}\frac{\partial \sigma_t^2}{\partial \omega}\left(\frac{\varepsilon_t^2-\sigma_t^2}{\sigma_t^2}\right),$$

$$\frac{\partial^2 l_t(y_t;b,\omega)}{\partial \omega \partial \omega'} = \left(\frac{\varepsilon_t^2-\sigma_t^2}{\sigma_t^2}\right)\frac{\partial}{\partial \omega'}\left[\frac{1}{2\sigma_t^2}\frac{\partial \sigma_t^2}{\partial \omega}\right] - \frac{1}{2\sigma_t^4}\frac{\partial \sigma_t^2}{\partial \omega}\frac{\partial \sigma_t^2}{\partial \omega'}\frac{\varepsilon_t^2}{\sigma_t^2},$$

where

$$\frac{\partial \sigma_t^2}{\partial \omega} = s_t + \sum_{i=1}^{p} b_i \frac{\partial \sigma_{t-1}^2}{\partial \omega}.$$

The first and second derivatives of the log-likelihood with respect to the mean parameter vector are:

$$\frac{\partial l_t(y_t;b,\omega)}{\partial b} = \varepsilon_t x_t \sigma_t^{-2} + \frac{1}{2}\sigma_t^2 \frac{\partial \sigma_t^2}{\partial b}\left(\frac{\varepsilon_t^2-\sigma_t^2}{\sigma_t^2}\right),$$

$$\frac{\partial^2 l_t(y_t;b,\omega)}{\partial b \partial b'} =$$

$$= -\sigma_t^{-2}x_t x_t' - \frac{1}{2}\sigma_t^{-4}\frac{\partial \sigma_t^2}{\partial b}\frac{\partial \sigma_t^2}{\partial b'}\left(\frac{\varepsilon_t^2}{\sigma_t^2}\right) - 2\sigma_t^{-4}\varepsilon_t x_t \frac{\partial \sigma_t^2}{\partial b} + \left(\frac{\varepsilon_t^2-\sigma_t^2}{\sigma_t^2}\right)\frac{\partial}{\partial b'}\left[\frac{1}{2}\sigma_t^{-2}\frac{\partial \sigma_t^2}{\partial b}\right],$$

where

$$\frac{\partial \sigma_t^2}{\partial b} = -2\sum_{i=1}^{q} a_i x_{t-i}\varepsilon_{t-i} + \sum_{j=1}^{p} b_j \frac{\partial \sigma_{t-j}^2}{\partial b}.$$

The information matrix corresponding to ω is given as:

$$I_{\omega\omega} = \frac{-1}{T}\sum_{t=1}^{T}\left(E\left(\frac{\partial^2 l_t(y_t;b,\omega)}{\partial \omega \partial \omega'}\right)\right) = \frac{1}{2T}\sum_{t=1}^{T}\left(\sigma_t^{-4}\frac{\partial \sigma_t^2}{\partial \omega}\frac{\sigma_t^2}{\partial \omega'}\right).$$

The information matrix corresponding to b is given as:

$$I_{bb} = \frac{-1}{T}\sum_{t=1}^{T}\left(E\left(\frac{\partial^2 l_t(y_t;b,\omega)}{\partial b \partial b'}\right)\right) = \frac{1}{T}\sum_{t=1}^{T}\left(\sigma_t^{-2}x_t x_t' + 2\sigma_t^{-4}\sum_{i=1}^{q} a_i^2 \varepsilon_{t-i}' x_{t-i} x_{t-i}' \varepsilon_{t-i} + \frac{1}{2}\sum_{j=1}^{p} b_j^2\left(\frac{\partial \sigma_{t-j}^2}{\partial b}\right)^2\right).$$

The elements in the off-diagonal block of the information matrix are zero, i.e.,

$$I_{\omega b} = \frac{-1}{T}\sum_{t=1}^{T}\left(E\left(\frac{\partial^2 l_t(y_t;b,\omega)}{\partial \omega \partial b'}\right)\right) = 0.$$

So, ω can be estimated without loss of asymptotic efficiency based on a consistent estimate of b and vice versa. At this point, it should be noticed that although the block diagonality holds for models as the GARCH, NARCH and Log-GARCH models, it does

not hold for asymmetric models, i.e. the EGARCH model, and for the ARCH in mean models. In such cases, the parameters have to be estimated jointly.

Even in the case of the symmetric GARCH(p,q) model with normally distributed innovations, we have to solve a set of $\breve{\theta} = k + p + q + 1$ non-linear equations in (4.4). Numerical techniques are used in order to estimate the vector of parameters ψ.

4.2 Numerical Estimation Algorithms

The problem faced in non-linear estimation, as in the case of the ARCH models, is that there are no closed form solutions. So, an iterative method has to be applied to obtain a solution. Iterative optimization algorithms work by taking an initial set of values for the parameters, say $\psi^{(0)}$, then performing calculations based on these values to obtain a better set of parameters values $\psi^{(1)}$. This process is repeated until the likelihood function, in equation (4.2), no longer improves between iterations. If $\psi^{(0)}$ is a trial value of the estimate, then expanding $L_T(\{y_t\};\psi)/\partial\psi$ and retaining only the first power of $\psi - \psi^{(0)}$, we obtain

$$\frac{\partial L_T}{\partial \psi} \approx \frac{\partial L_T}{\partial \psi^{(0)}} + \left(\psi - \psi^{(0)}\right)\frac{\partial^2 L_T}{\partial \psi^{(0)} \partial \psi'^{(0)}} .$$

At the maximum, $L_T/\partial\psi$ should equal zero. Rearranging terms, the correction for the initial value, $\psi^{(0)}$, obtained is

$$\left(\psi - \psi^{(0)}\right) = -\frac{\partial L_T}{\partial \psi^{(0)}}\left(\frac{\partial^2 L_T}{\partial \psi^{(0)} \partial \psi'^{(0)}}\right)^{-1} . \qquad (4.6)$$

Let $\psi^{(i)}$ denote the parameter estimates after the i^{th} iteration. Based on (4.6) the Newton-Raphson algorithm computes $\psi^{(i+1)}$ as:

$$\psi^{(i+1)} = \psi^{(i)} - \left(\frac{\partial^2 L_T^{(i)}}{\partial \psi \partial \psi'}\right)^{-1} \frac{\partial L_T^{(i)}}{\partial \psi} . \qquad (4.7)$$

The scoring algorithm is a method closely related to the Newton-Raphson algorithm and was applied by Engle (1982) to estimate the parameters of the ARCH(p) model. The difference between the Newton-Raphson method and the method of scoring is that the

former depends on observed second derivatives, while the latter depends on the expected values of the second derivatives. So, the scoring algorithm computes $\psi^{(i+1)}$ as:

$$\psi^{(i+1)} = \psi^{(i)} + E\left(\frac{\partial^2 L_T^{(i)}}{\partial\psi\partial\psi'}\right)^{-1} \frac{\partial L_T^{(i)}}{\partial\psi}. \tag{4.8}$$

An alternative procedure suggested by Berndt et al. (1974), which uses first derivatives only, is the Berndt, Hall, Hall and Hausman (BHHH) algorithm. The BHHH algorithm is similar to the Newton-Raphson algorithm, but, instead of the Hessian (second derivative of the log likelihood function with respect to the vector of unknown parameters), it is based on an approximation formed by the sum of the outer product of the gradient vectors for the contribution of each observation to the objective function. This approximation is asymptotically equivalent to the actual Hessian when evaluated at the parameter values, which maximize the function. The BHHH algorithm computes $\psi^{(i+1)}$ as:

$$\psi^{(i+1)} = \psi^{(i)} + \left(\sum_{t=1}^{T} \frac{\partial l_t^{(i)}}{\partial\psi} \frac{\partial l_t^{(i)}}{\partial\psi'}\right)^{-1} \frac{\partial L_T^{(i)}}{\partial\psi}. \tag{4.9}$$

When the outer product is near singular, a ridge correction may be used in order to handle numerical problems and improve the convergence rate. Marquardt (1963) modified the BHHH algorithm by adding a correction matrix to the sum of the outer product of the gradient vectors. The Marquardt updating algorithm is computed as:

$$\psi^{(i+1)} = \psi^{(i)} + \left(\sum_{t=1}^{T} \frac{\partial l_t^{(i)}}{\partial\psi} \frac{\partial l_t^{(i)}}{\partial\psi'} - aI\right)^{-1} \frac{\partial L_T^{(i)}}{\partial\psi}, \tag{4.10}$$

where I is the identity matrix and a is a positive number chosen by the algorithm. The effect of this modification is to push the parameter estimates in the direction of the gradient vector. The idea is that when we are far from the maximum, the local quadratic approximation to the function may be a poor guide to its overall shape, so it may be better off simply following the gradient. The correction may provide a better performance at locations far from the optimum, and allows for computation of the direction vector in cases where the Hessian is near singular.

4.3 Maximum Likelihood Estimation under Non-Normality

As already mentioned, an attractive feature of the ARCH process is that even though the conditional distribution of the innovations is normal, the unconditional distribution has thicker tails than the normal one. However, the degree of leptokurtosis induced by the ARCH process often does not capture all of the leptokurtosis present in high frequency speculative prices. Thus, there is a fair amount of evidence that the conditional distribution of ε_t is non-normal as well.

To circumvent this problem, Bollerslev (1987) proposed using the standardized t distribution with $v > 2$ degrees of freedom:

$$f(z_t; v) = \frac{\Gamma\left((v+1)/2\right)}{\Gamma(v/2)\sqrt{\pi(v-2)}}\left(1+\frac{z_t^2}{v-2}\right)^{-\frac{v+1}{2}}, \quad v > 2, \tag{4.11}$$

where $\Gamma(.)$ is the gamma function. The degrees of freedom are regarded as parameter to be estimated, $w = (v)$. The t distribution is symmetric around zero and for $v > 4$ the conditional kurtosis equals $3(v-2)(v-4)^{-1}$, which exceeds the normal value of three, but for $v \to \infty$, (4.11) converges to (4.5), the standard normal distribution.

Nelson (1991) suggested the use of the generalized error distribution, or GED[4]:

$$f(z_t; v) = \frac{v\exp\left(-0.5\left|z_t/\lambda\right|^v\right)}{\lambda 2^{\left(1+1/v\right)}\Gamma\left(v^{-1}\right)}, \quad v > 0, \tag{4.12}$$

where v is the tail-thickness parameter and $\lambda \equiv \sqrt{2^{-2/v}\Gamma\left(v^{-1}\right)\Gamma\left(3v^{-1}\right)}$. (For more details on the GED, see Harvey (1981) and Box and Tiao (1973)). When $v = 2$, z_t is standard normally distributed and so (4.12) reduces to (4.5). For $v < 2$, the distribution of z_t has thicker tails than the normal distribution (e.g., for $v = 1$, z_t has a double exponential distribution) while for $v > 2$, the distribution of z_t has thinner tails than the normal distribution (e.g., for $v = \infty$, z_t has a uniform distribution on the interval $(-\sqrt{3}, \sqrt{3})$).

The densities presented above account for fat tails but they are symmetric. Lee and Tse (1991) suggested that not only the conditional distribution of innovations may be

[4] The GED sometimes referred as the exponential power distribution.

leptokurtotic, but also asymmetric. Allowing for skewness may be important in modeling interest rates as they are lower bounded by zero and may therefore be skewed. To allow for both skewness and leptokurtosis, they used a Gram Charlier type distribution (see Kendall and Stuart (1969), p.157) with density function given by:

$$f(z_t;v,g)=\widehat{f}(z_t)\left(1+\frac{v}{6}H_3(z_t)+\frac{g}{24}H_4(z_t)\right), \tag{4.13}$$

where $\widehat{f}(.)$ is the standard normal density function, and $H_3(z_t)\equiv z_t^3-3z_t$ and $H_4(z_t)\equiv z_t^4-6z_t^2+3$ are the Hermite polynomials. The quantities v and g are the measures of skewness and kurtosis, respectively. Jondeau and Rockinger (2001) examined the properties of the Gram Charlier conditional density function and estimated ARCH models with a Gram Charlier density function for a set of exchange rate series.

Bollerslev et al. (1994) applied the generalized t distribution (McDonald and Newey (1988)):

$$f(z_t;v,g)=\frac{v}{2\sigma_t b g^{1/v}B(v^{-1},g)\left(1+\left(|\varepsilon_t|^v/gb^v\sigma_t^v\right)\right)^{g+1/v}},\ v>0,\ g>0 \text{ and } vg>2, \tag{4.14}$$

where $B(v^{-1},g)\equiv\Gamma(v^{-1})\Gamma(g)/\Gamma(v^{-1}+g)$ is the beta function and $b\equiv\sqrt{\Gamma(v^{-1})\Gamma(g)/\Gamma(3v^{-1})\Gamma(g-2v^{-1})}$. The generalized t distribution has the advantage that nests both (4.11) and (4.12). For $v=2$ and $g=0.5$ times the degrees of freedom, (4.14) is set to the t distribution, and for $v=\infty$, the GED is obtained. Moreover, the two shape parameters v and g allow for fitting both the tails and the central part of the conditional distribution.

Lambert and Laurent (2000, 2001) extended the skewed Student t density proposed by Fernandez and Steel (1998) to the ARCH framework, in the following density function:

$$f(z_t;v,g)=\frac{\Gamma((v+1)/2)}{\Gamma(v/2)\sqrt{\pi(v-2)}}\left(\frac{2s}{g+g^{-1}}\right)\left(1+\frac{sz_t+m}{v-2}g^{-II_t}\right)^{-\frac{v+1}{2}},\quad v>2, \tag{4.15}$$

where g is the asymmetry parameter, v denotes the degrees of freedom of the distribution, $\Gamma(.)$ is the gamma function, $II_t=1$ if $z_t\geq -ms^{-1}$, and $II_t=-1$ otherwise, $m=\Gamma((v-1)/2)\sqrt{(v-2)}\left(\Gamma(v/2)\sqrt{\pi}\right)^{-1}\left(g-g^{-1}\right)$ and $s=\sqrt{g^2+g^{-2}-m^2-1}$.

Vries (1991) noted that the unconditional distribution of variaties from an ARCH process can be stable and that under suitable conditions the conditional distribution is stable as well. Stable Paretian conditional distributions have been introduced in ARCH models by Liu and Brorsen (1995), Mittnik et al. (1999), and Panorska et al. (1995). As the stable Paretian distribution does not have an analytical expression for its density function, it is expressed by its characteristic function:

$$\varphi(t,a,\beta,\sigma,\mu)=\exp\left(i\mu t-|\sigma t|^{a}\left(1-i\beta\frac{t}{|t|}\omega(|t|,a)\right)\right), \tag{4.16}$$

where $0<a\le 2$ is the characteristic exponent, $-1\le\beta\le 1$ is the skewness parameter, $\sigma>0$ is the scale parameter, $\mu\in\Re$ is the location parameter, and

$$\omega(|t|,a)=\begin{cases}\tan\dfrac{\pi a}{2}, & \mathrm{a}\neq 1\\ -\dfrac{2}{\pi}\log|t|, & \mathrm{a}=1.\end{cases}$$

The standardized innovations, z_t, are assumed as independently, identically stable Pareto distributed random variables with zero location parameter and unit scale parameter. The way that GARCH models are built imposes limits on the heaviness of the tails of their unconditional distribution. Given that a wide range of financial data exhibit remarkable fat tails, this assumption represents a major shortcoming of GARCH models in financial time series analysis. Stable Paretian conditional distributions have been employed in a number of studies, such as Mittnik et al. (1998a, 1998b) and Mittnik and Paolella (2001). Tsionas (1999) established a framework for Monte Carlo posterior inference in models with stable distributed errors by combining a Gibbs sampler with Metropolis independence chains and representing the symmetric stable variates as normal scale mixtures. Mittnik et al. (2002) and Panorska et al. (1995) derived conditions for strict stationarity of GARCH and APARCH models with stable Paretian conditional distributions. Vries (1991) provided relationships between ARCH and stable processes. Tsionas (2002) compared a stable Paretian model with ARCH errors with a stable Paretian model with stochastic volatility. The Randomized GARCH model with stable Paretian innovations totally skewed to the right and with $0<a<1$ was studied by Nowicka-Zagrajek and Weron (2001). They derived the unconditional distributions and analyzed the dependence structure by means of the codifference. It turns out that R-GARCH models with conditional variance dependent on the past can have very heavy

tails. The class is very flexible as it includes GARCH models and Vries process (1991) as special cases.

Hansen (1994) suggested an approach that allows not only the conditional variance to be time varying but also the higher moments of conditional distribution such as skewness and kurtosis. He suggested the autoregressive conditional density, or the ARCD, model, where the density function, $f(z_t; w)$, is presented as:

$$f(z_t; w_t \mid I_{t-1}). \tag{4.17}$$

The parameter vector of the conditional density function in (4.17) is assumed to be a function of the current information set, I_{t-1}.

Other distributions, that have been employed, include the normal Poisson mixture distribution (Brorsen and Yang (1994), Drost et al. (1998), Jorion (1988), Lin and Yeh (2000), and Vlaar and Palm (1993)), the normal lognormal mixture (Hsieh (1989)), and serially dependent mixture of normally distributed variables (Cai (1994)) or student t distributed variables (Hamilton and Susmel (1994))[5].

4.4 Quasi-Maximum Likelihood Estimation

The assumption of normally distributed standardized innovations is often violated by the data. This has motivated the use of alternative distributional assumptions, presented in the previous section. Alternatively, the MLE based on the normal density may be given a quasi-maximum likelihood interpretation. Bollerslev and Wooldridge (1992), based on Weiss (1986) and Pagan and Sabau (1987), showed that the maximization of the normal log-likelihood function can provide consistent estimates of the parameter vector θ even when the distribution of z_t in non-normal, provided that:

$$E(z_t \mid I_{t-1}) = 0$$
$$E(z_t^2 \mid I_{t-1}) = 1.$$

This estimator is, however, inefficient with the degree of inefficiency increasing with the degree of departure from normality. So, the standard errors of the parameters have to be adjusted. Let $\hat{\theta}$ be the estimate that maximizes the normal log-likelihood function, in

[5] Cai (1994) and Hamilton and Susmel (1994) used the mixtures to estimate the class of regime switching ARCH models, presented in section 2.1.

equation (4.2), based on the normal density function in (4.5), and let θ_0 be the true value. Then, even when z_t is non-normal, under certain regularity conditions:

$$\sqrt{T}\left(\hat{\theta}-\theta_0\right)\xrightarrow{D} N\left(0, A^{-1}BA^{-1}\right), \tag{4.17}$$

where

$$A \equiv \underset{T\to\infty}{p\lim} T^{-1}\sum_{t=1}^{T} E\left(\frac{-\partial^2 l_t(\theta_0)}{\partial\theta\partial\theta'}\right),$$

$$B \equiv \underset{T\to\infty}{p\lim} T^{-1}\sum_{t=1}^{T} E\left(\frac{\partial l_t(\theta_0)}{\partial\theta}\right)\left(\frac{\partial l_t(\theta_0)}{\partial\theta}\right)',$$

for l_t denoting the correctly specified log-likelihood function. The matrices A and B can be consistently estimated by:

$$\hat{A} = -T^{-1}\sum_{t=1}^{T} E\left(\frac{\partial^2 l_t(\hat{\theta})}{\partial\theta\partial\theta'} \mid I_{t-1}\right),$$

$$\hat{B} = T^{-1}\sum_{t=1}^{T} E\left(\left(\frac{\partial l_t(\hat{\theta})}{\partial\theta}\right)\left(\frac{\partial l_t(\hat{\theta})}{\partial\theta'}\right) \mid I_{t-1}\right),$$

where l_t is the incorrectly specified log-likelihood function under the assumption of normal density function. Thus, standard errors for $\hat{\theta}$ that are robust to misspecification of the family of densities can be obtained from the square root of diagonal elements of:

$$T^{-1}\hat{A}^{-1}\hat{B}\hat{A}^{-1}.$$

Recall that if the model is correctly specified and the data are in fact generated by the normal density function, then $A = B$, and, hence, the variance covariance matrix, $T^{-1}\hat{A}^{-1}\hat{B}\hat{A}^{-1}$, reduces to the usual asymptotic variance covariance matrix for maximum likelihood estimation:

$$T^{-1}\hat{A}^{-1}.$$

For symmetric departures from normality, the quasi-maximum likelihood estimation is generally close to the exact MLE. But, for non-symmetric distributions, Engle and González-Rivera (1991), showed that the loss in efficiency may be quite high (Bai and Ng (2001) proposed a procedure for testing conditional symmetry.). In such a case, other methods of estimation should be considered. Lumsdaine (1991, 1996) and Lee and Hansen (1991, 1994) established the consistency and asymptotic normality of the quasi-

maximum likelihood estimators of the IGARCH(1,1) model. Lee (1991) extended the asymptotic properties to the IGARCH(1,1) in Mean model, Berkes et al. (2003) and Berkes and Horváth (2003) studied the asymptotic properties of the quasi-maximum likelihood estimators of the GARCH(p,q) model under a set of weaker conditions, and Baille et al. (1996) showed that the quasi-maximum likelihood estimators of the FIGARCH(1,d,0) model are both consistent and asymptotically normally distributed.

4.5 Other Estimating Methods

Other estimation methods, except for MLE, have been appeared in the ARCH literature. Harvey et al. (1992) presented the unobserved components structural ARCH, or STARCH, model and proposed an estimation method based on the Kalman filter. These are state space models or factor models in which the innovation is composed of several sources of error where each of the error sources has a heteroscedastic specification of the ARCH form. Since the error components cannot be separately observed given the past observations, the independent variables in the variance equations are not measurable with respect to the available information set, which complicates inference procedures.

Pagan and Hong (1991) applied a nonparametric Kernel estimate of the expected value of squared innovations. Pagan and Schwert (1990) used a collection of nonparametric estimation methods, including Kernels, Fourier series and two-stage least squares regressions. They found that the non-parametric methods did good job in-sample forecasts though the parametric models yielded superior out-of-sample forecasts. Gouriéroux and Monfort (1992) also proposed a nonparametric estimation method in order to estimate the GQTARCH model in equation (2.28). Bühlmann and McNeil (2002) proposed a nonparametric estimation iterative algorithm, that requires neither the specification of the conditional variance functional form nor that of the conditional density function, and showed that their algorithm gives more precise estimates of the volatility in the presence of departures from the assumed ARCH specification.

Engle and González-Rivera (1991), Engle and Ng (1993), Gallant and Tauchen (1989), Gallant et al. (1991), Gallant et al. (1993) among others, combined parametric specifications for the conditional variance with a nonparametric estimate of the conditional density function. In a Monte Carlo study, Engle and González-Rivera (1991)

found that their semi-parametric method could improve the efficiency of the parameter estimates up to 50 per cent over the QMLE, particularly when the density was highly non-normal and skewed, but it did not seem to capture the total potential gain in efficiency.

Another attractive way to estimate ARCH models without assuming normality is to apply the generalized method of moments (GMM) approach. (For details, see Bates and White (1988), Ferson (1989), Mark (1988), Rich et al. (1991), Simon (1989)). Let us, for example, represent the GARCH(p,q) model as $\sigma_t^2 = \omega' s_t$, where $\omega' = \left(a_0, a_1, ..., a_q, b_1, ..., b_p\right)$ and $s_t = \left(1, \varepsilon_{t-1}^2, ..., \varepsilon_{t-q}^2, \sigma_{t-1}^2, ..., \sigma_{t-p}^2\right)$. Under the assumption of:

$$E\left(\left(y_t - x_t' b\right)x_t\right) = 0$$
$$E\left(\left(\varepsilon_t^2 - \sigma_t^2\right)s_t\right) = 0,$$

the parameters could be estimated by GMM by choosing the vector $\theta' = \left(b', \omega'\right)$ so as to minimize:

$$\left(g\left(\theta; I_{t-1}\right)\right)' \hat{S} \left(g\left(\theta; I_{t-1}\right)\right),$$

where

$$g\left(\theta; I_{t-1}\right) = \begin{bmatrix} T^{-1}\sum_{t=1}^{T}\left(y_t - x_t' b\right)x_t \\ T^{-1}\sum_{t=1}^{T}\left(\left(y_t - x_t' b\right)^2 - \omega' s_t\right)s_t \end{bmatrix},$$

and the matrix $\hat{S}$ can be constructed by any of the methods that have been considered in the GMM literature.

Geweke (1988a,b, 1989) argued that a Bayesian approach, rather than the classical one, might be more suitable for estimating ARCH models due to the distinct features of these models. In order to ensure positivity of the conditional variance, some inequality restrictions should be imposed. Although difficult to impose such restrictions in the classical approach, under the Bayesian framework, diffuse priors can incorporate these inequalities. Also, as the main interest in not in the individual parameters but rather in the conditional variance itself, in the Bayesian framework exact posterior distributions of the conditional variance can be obtained.

Giraitis and Robinson (2000) estimated the parameters of the GARCH process using the Whittle estimation technique and demonstrated that the Whittle estimator is

strongly consistent and asymptotically normal, provided the GARCH process has finite 8th moment marginal distribution. Whittle (1953) proposed an estimation technique that works in the spectral domain of the process[6]. Moreover, Mikosch and Straumann (2002) showed that the Whittle estimator is consistent as long as the 4th moment is finite and inconsistent when the 4th moment is infinite. Thus, as noted by Mikosch and Straumann, the Whittle estimator for GARCH processes is unreliable as the ARCH models are applied in heavy-tailed data, sometimes without finite 5th, 4th, or even 3rd moments.

Hall and Yao (2003) showed that for heavy tailed innovations, the asymptotic distribution of quasi-maximum likelihood parameter estimators is non-normal and suggested percentile-t subsample bootstrap approximations to estimator distributions.

5. Multivariate ARCH Models

All the ARCH models that have been discussed are univariate. However, assets and markets affect each other not only in terms of expected returns but also in terms of volatility. Thus, the accurate estimation of time-varying covariances between asset returns has been crucial for asset pricing and risk management. The generalization of univariate models to a multivariate context leads to a straightforward application of ARCH models to portfolio selection and asset pricing theory.

Let the $(n\times 1)$ vector $\{\mathbf{y}_t\}$ refer to the multivariate discrete time real-valued stochastic process to be predicted, where $E_{t-1}(\mathbf{y}_t)\equiv\boldsymbol{\mu}_t$ denotes the conditional mean. The innovation process for the conditional mean $\boldsymbol{\varepsilon}_t\equiv\mathbf{y}_t-\boldsymbol{\mu}_t$ has an $(n\times n)$ conditional covariance matrix $V_{t-1}(\mathbf{y}_t)\equiv\mathbf{H}_t$. For a system of n regression equations, the natural extension of (2.2) to a multivariate framework could be presented as:

$$\begin{aligned}\mathbf{y}_t &= \mathbf{B}'\mathbf{x}_t+\boldsymbol{\varepsilon}_t\\ \boldsymbol{\varepsilon}_t \mid I_{t-1} &\sim f[0,\mathbf{H}_t]\\ \mathbf{H}_t &= g(\mathbf{H}_{t-1},\mathbf{H}_{t-2},...,\boldsymbol{\varepsilon}_{t-1},\boldsymbol{\varepsilon}_{t-2},...),\end{aligned} \tag{5.1}$$

where $\mathbf{B}$ is a $k\times n$ matrix of unknown parameters, $\mathbf{x}_t$ a $k\times 1$ vector of endogenous and exogenous explanatory variables included in the available information set, I_{t-1}, $f(.)$ the

[6] For further details about the Whittle estimation technique for ARMA processes see Brockwell and Davis (1991).

conditional multivariate density function of innovation process and $g(.)$ a function of the lagged conditional covariance matrices and innovation process.

The natural multivariate extension of the GARCH(p,q) model in equation (2.6) is:

$$\mathbf{H}_t = \mathbf{A_0}\mathbf{A_0}' + \sum_{i=1}^{q}\left(\mathbf{A}_i\boldsymbol{\varepsilon}_{t-i}\boldsymbol{\varepsilon}_{t-i}'\mathbf{A}_i'\right) + \sum_{j=1}^{p}\left(\mathbf{B}_j\mathbf{H}_{t-j}\mathbf{B}_j'\right), \tag{5.2}$$

where $\mathbf{A_0}$ is a lower triangular matrix with $\left(n(n+1)/2\right)$ parameters and, $\mathbf{A}_i$ and $\mathbf{B}_j$ denote $(n\times n)$ matrices with n^2 parameters each. Engle and Kroner (1995), based on an earlier work of Baba et al. (1990), proposed model (5.2) to which they referred as the BEKK model. This parameterization guarantees that $\mathbf{H}_t$ is positive definite and requires the estimation of $\left(n(n+1)/2\right)+n^2(q+p)$ parameters. For example, for $n=3$, the multivariate GARCH(1,1) model contains 24 parameters for estimation. Lee (1999b) investigated the output-inflation variability tradeoff using the bivariate BEKK model. Recently, Moschini and Myers (2002), in order to estimate time-varying optimal hedge ratios in commodity markets, modified the BEKK model of (5.2) in the form:

$$\mathbf{H}_t = \boldsymbol{\Gamma}_t'\left(\mathbf{A_0}\mathbf{A_0}' + \sum_{i=1}^{q}\left(\mathbf{A}_i\boldsymbol{\varepsilon}_{t-i}\boldsymbol{\varepsilon}_{t-i}'\mathbf{A}_i'\right) + \sum_{j=1}^{p}\left(\mathbf{B}_j\mathbf{H}_{t-j}\mathbf{B}_j'\right)\right)\boldsymbol{\Gamma}_t .$$

As Moschini and Myers noted, the covariance matrix is positive definite as long as $\boldsymbol{\Gamma}_t$ is a positive definite matrix.

A simpler expression of $\mathbf{H}_t$ can be obtained through the use of the $vech(.)$ operator that stacks the lower portion of a $(n\times n)$ matrix as an $\left(n(n+1)/2\right)\times 1$ vector. So, the equation (5.2) is rewritten as:

$$vech(\mathbf{H}_t) = vech(\mathbf{A_0}\mathbf{A_0}') + \sum_{i=1}^{q}\left(\tilde{\mathbf{A}}_i\, vech(\boldsymbol{\varepsilon}_{t-i}\boldsymbol{\varepsilon}_{t-i}')\right) + \sum_{j=1}^{p}\left(\tilde{\mathbf{B}}_j\, vech(\mathbf{H}_{t-j})\right), \tag{5.3}$$

where $\tilde{\mathbf{A}}_i$ and $\tilde{\mathbf{B}}_j$ are parameter matrices of dimension $\left(n(n+1)/2\times n(n+1)/2\right)$. Engle et al. (1986) published the first paper on multivariate ARCH models applying the multivariate ARCH(2) model. However, in the multivariate expression of the GARCH(p,q) model, serious problems arise: i) the model might not yield a positive definite covariance matrix unless nonlinear inequality restrictions are imposed, and ii) the number of parameters has to be estimated is $\left(n(n+1)/2\right)\left(1+\left(n(n+1)/2\right)(q+p)\right)$, a very large

number even for low dimensions of n. For example, for $n=3$, the multivariate GARCH(1,1) model contains 78 parameters for estimation.

A number of models, considered in the financial literature, have dealt with imposing constraints in multivariate GARCH models in order to reduce the number of parameters that should be estimated. These constraints have to be compatible with a positive definite conditional covariance matrix and must lead to tractable estimation procedures. Bollerslev et al. (1988) proposed the diagonal multivariate GARCH(p,q) model where the $\tilde{\mathbf{A}}_i$ and $\tilde{\mathbf{B}}_j$ matrices are supposed to be diagonal. Thus, the number of parameters is reduced to $\left(n(n+1)/2\right)(1+q+p)$. So, for example, for $n=3$, the diagonal GARCH(1,1) model requires the estimation of 18 parameters. Bollerslev et al. (1988) used this model for analyzing returns on bills, bonds and stocks, while Baillie and Myers (1991), Bera et al. (1991) and Myers (1991) estimated hedge ratios in commodity markets. Ding and Engle (2001) gave sufficient conditions for the diagonal multivariate GARCH(1,1) model to be positive definite and proposed four models, which are nested to the multivariate diagonal multivariate GARCH(1,1) model.

A special case of the BEKK model, for $p=q=1$, is the factor GARCH model first proposed in Engle (1987). The factor GARCH(1,1) model was constructed to overcome the problem of estimating a vast number of parameters, while retaining the benefits of positive definiteness. The model has the form:

$$\mathbf{H}_t = \mathbf{A_0}\mathbf{A_0}' + \boldsymbol{\lambda}\boldsymbol{\lambda}'\left(\alpha\left(\mathbf{w}'\boldsymbol{\varepsilon}_{\mathbf{t-1}}\right)^2 + \beta\mathbf{w}'\mathbf{H}_{t-1}\mathbf{w}\right), \tag{5.4}$$

where α and β are scalars, $\boldsymbol{\lambda}$ and $\mathbf{w}$ are $(n\times 1)$ vectors. The vector $\mathbf{w}$ can be considered as a vector of portfolio weights and it is convenient to restrict in the case $\boldsymbol{\iota}'\mathbf{w}=1$, where $\boldsymbol{\iota}$ is a vector of ones. This model is a special case of the BEKK model where the matrices $\mathbf{A}_1$ and $\mathbf{B}_1$ have rank 1: $\mathbf{A}_1=\sqrt{\alpha}\mathbf{w}\boldsymbol{\lambda}'$ and $\mathbf{B}_1=\sqrt{\beta}\mathbf{w}\boldsymbol{\lambda}'$. The number of parameters is $\left(n^2+5n\right)/2+2$. So for example, for $n=3$ we have to estimate 14 parameters. The model can be extended to allow for K factors and a higher order GARCH structure. So, the K factor GARCH(p,q) model is represented by

$$\mathbf{H}_t = \mathbf{A_0}\mathbf{A_0}' + \sum_{k=1}^{K}\boldsymbol{\lambda}_k\boldsymbol{\lambda}_k'\left(\sum_{i=1}^{q}\alpha_{k,i}\mathbf{w}_k'\boldsymbol{\varepsilon}_{t-i}\boldsymbol{\varepsilon}_{t-i}'\mathbf{w}_k + \sum_{j=1}^{p}\beta_{k,j}\mathbf{w}_k'\mathbf{H}_{t-j}\mathbf{w}_k\right), \tag{5.5}$$

and has $K(2n+p+q)+n(n+1)/2$ free parameters. Engle et al. (1990b) and Ng et al. (1992) applied factor GARCH models on treasury bills and stock returns. Diebold and

Nerlove (1989), Harvey et al. (1992), King et al. (1994) and Alexander (2000) proposed latent factor GARCH models, based on the assumption that only a few factors influence the conditional variances and covariances of asset returns, which are not functions of the information set.

The constant conditional correlation model, introduced by Bollerslev (1990), is a popular method to model multivariate GARCH models, where univariate GARCH models are estimated for each asset and then the correlation matrix is estimated. The time-varying conditional covariances are parameterized to be proportional to the product of the corresponding conditional standard deviations. This assumption greatly simplifies the estimation of the model and reduces the computational cost. Let us assume that the covariance matrix can be decomposed thus $\mathbf{H}_t = \mathbf{\Sigma}_t^{1/2}\mathbf{C}_t\mathbf{\Sigma}_t^{1/2}$, where $\mathbf{\Sigma}_t$ is the diagonal matrix with the conditional variances along the diagonal and $\mathbf{C}_t$ is the matrix of conditional correlations. The constant conditional correlation model assumes that the matrix of conditional correlations is time invariant, so that the temporal variation of $\mathbf{H}_t$ can be determined solely by the conditional variances:

$$\mathbf{H}_t = \mathbf{\Sigma}_t^{1/2}\mathbf{C}\mathbf{\Sigma}_t^{1/2}. \tag{5.6}$$

$\mathbf{H}_t$ is positive definite if $\mathbf{C}$ is positive definite and the conditional variances are positive. The number of parameters reduces to $\left(n(n-1)/2\right)+n(1+q+p)$. So, for $n=3$ the constant conditional correlation GARCH(1,1) model requires the estimation of 12 parameters. Several authors have considered this representation, e.g. Baillie and Bollerslev (1990), Brown and Ligeralde (1990), Cecchetti et al. (1988), Fornari et al. (2002), Kim (2000), Kroner and Claessens (1991), Kroner and Lastrapes (1991), Kroner and Sultan (1991,1993), Lien and Tse (1998) and Park and Switzer (1995).

However, recent studies have considered test statistics, which reject the constancy of conditional correlation. Bera and Kim (1996), who proposed the Information Matrix test, were led to the rejection of a constant correlation hypothesis for USA, European and Japan stock markets, while Tse (2000), who derived a Lagrange Multiplier test for the conditional correlation stability hypothesis, rejected the hypothesis for Asian stock markets. Tsui and Yu (1999), adopting the Information Matrix test, examined the China stock market and found that the constant conditional correlation hypothesis is not supported. Longin and Solnik (1995) rejected the hypothesis of constant conditional correlation in international equity returns against three alternative sources of variability of

the correlation such as a time trend, the presence of threshold and asymmetry and the influence of information variables.

As the hypothesis of constancy of correlation was rejected in a number of papers, Engle (2000) and Engle and Sheppard (2001) introduced a new form of multivariate ARCH model, the Dynamic Conditional Correlation GARCH, or DCC-GARCH(C,M), model. The model is estimated in two steps. The first is a series of univariate GARCH estimates. The second step, using the residuals resulting for the first stage, evaluates the conditional correlation estimator. The success of the DCC-GARCH model depends on the estimability of extremely large time varying covariance matrices. Engle proposed to use the decomposed covariance matrix $\mathbf{H}_t = \mathbf{\Sigma}_t^{1/2}\mathbf{C}_t\mathbf{\Sigma}_t^{1/2}$ and suggested a time varying correlation matrix of the following form:

$$\mathbf{C}_t = \mathbf{Q}_t^{*-1/2}\mathbf{Q}_t\mathbf{Q}_t^{*-1/2}. \tag{5.7}$$

The conditional variances, $\sigma_{k,t}^2$, are estimated as univariate GARCH(p_k, q_k) models, allowing for different lag lengths for each series $k = 1,2,...,n$,

$$\sigma_{k,t}^2 = a_{k,0} + \sum_{i=1}^{q_k}\left(a_{k,i}\varepsilon_{k,t-i}^2\right) + \sum_{j=1}^{p_k}\left(b_{k,j}\sigma_{k,t-j}^2\right). \tag{5.8}$$

The correlation matrix is computed using

$$\mathbf{Q}_t = \left(1 - \sum_{m=1}^{M} a_m - \sum_{c=1}^{C} b_c\right)\overline{\mathbf{Q}} + \sum_{m=1}^{M} a_m\left(\mathbf{z}_{t-m}\mathbf{z}'_{t-m}\right) + \sum_{c=1}^{C} b_c\mathbf{Q}_{t-c}, \tag{5.9}$$

where $\mathbf{z}_t$ are the residuals standardized by their conditional standard deviation, $\overline{\mathbf{Q}}$ is the unconditional covariance of the standardized residuals and $\mathbf{Q}_t^{*-1/2}$ is a diagonal matrix composed of the square roots of the diagonal elements of $\mathbf{Q}_t$. Engle and Sheppard (2001) proved the consistency and asymptotic normality of the two-step estimators as well as the positive definiteness of the covariance matrix. They have also proposed a test of the null hypothesis of constant correlation against an alternative of dynamic conditional correlation. Christodoulakis and Satchell (2002) considered an alternative extension of the constant conditional correlation model of Bollerslev (1990) and developed a bivariate ARCH model with time varying conditional variances and correlations, named Correlated ARCH, or CorrARCH, model.

The multivariate ARCH models, that have been presented, although simplifying the estimation and inference procedures, do not account for empirical regularities such

as the asymmetric effects. In to order to capture the "leverage effect" in a multivariate framework, Braun et al. (1995) introduced a bivariate version of the EGARCH model in equation (2.13). Sentana (1995), in the presentation of the quadratic GARCH model, applied a multivariate version of his model to U.K. stock returns. Kroner and Ng (1998), following Hentschel's (1995) approach, introduced a general multivariate GARCH model which nests the BEKK, diagonal, factor and constant conditional correlation GARCH models and their natural asymmetric extensions. Their model can be regarded as a multivariate extension of the GJR model in equation (2.13). Bekaert and Wu (1997), Ding and Engle (2001) and Tai (2001) have also modified multivariate ARCH models to accommodate asymmetric effects on conditional variances and covariances. Brunetti and Gilbert (1998), based on Bollerslev's (1990) parameterization, proposed the bivariate constant correlation FIGARCH model and Brunetti and Gilbert (2000) applied the model to the crude oil market.

6. Other Methods of Volatility Modeling

"*Stochastic volatility*" models (Barndorff-Nielsen et al. (2001), Chib et al. (1998), Ghysels et al. (1996), Harvey and Shephard (1993), Jacquier et al. (1994), Shephard (1996), Taylor (1994)), "*implied volatility*" models (Day and Lewis (1988), Latane and Rendleman (1976), Schmalensee and Trippi (1978)), "*historical volatility*" models (Beckers (1983), Garman and Klass (1980), Kunitomo (1992), Parkinson (1980), Rogers and Satchell (1991)) and "*realized volatility*" models are examples from the financial econometric literature of estimating volatility of asset returns.

A typical presentation of a stochastic volatility model can be given by

$$\begin{aligned} \varepsilon_t &= z_{1,t}\sigma e^{0.5\sigma_t} \\ \sigma_t^2 &= a\sigma_{t-1}^2 + z_{2,t} \\ z_{1,t} &\overset{i.i.d.}{\sim} f\left[E(z_t)=0, V(z_t)=1\right] \\ z_{2,t} &\overset{i.i.d.}{\sim} g\left[E(z_t)=0, V(z_t)=\sigma_{z_2}^2\right], \end{aligned} \tag{6.1}$$

where σ is a positive scale parameter, $|a|<1$, and the error terms $z_{1,t}$ and $z_{2,t}$ could be contemporaneously correlated. The additional error term, $z_{2,t}$, in the conditional variance equation makes the stochastic volatility model have no closed form solution. Hence, the estimation of the parameters is a quite difficult task. For this reason, stochastic volatility

models are not as popular as the ARCH processes. Jacquier et al. (1994) considered a Markov Chain Monte Carlo (MCMC) framework in order to estimate stochastic volatility models and Jacquier et al. (1999) extended the MCMC technique to allow for the leverage effect and fat tailed conditional errors. Nelson (1990b) was the first to show that the continuous time limit of an ARCH process, which is a stochastic difference equation, is a diffusion process with stochastic volatility (which is a stochastic differential equation). Duan (1996) extended Nelson's study.

Models based on the daily open, high, low and close asset prices, and exponential smoothing methods, such as the Riskmetrics method by J.P. Morgan, are procedures which are included to the historical volatility models.

Implied volatility is the instantaneous standard deviation of the return on the underlying asset, which would have to be input into a theoretical pricing model in order to yield a theoretical value identical to the price of the option in the marketplace, assuming all other inputs are known. Day and Lewis (1992) examined whether implied volatilities contain incremental information relative to the estimated volatility from ARCH models. Noh et al. (1994) compared the forecasting performance of ARCH and implied volatility models in the context of option pricing. Andersen et al. (2002) reviewed a systematically categorization of various ways of modeling volatility. Recently, Poon and Granger (2001) conducted a comparative review based on the forecasting performance of ARCH, implied volatility, and historical volatility models.

Although the presentation of the above methods of volatility estimation is beyond the scope of this paper, we briefly refer to the modeling of realized volatility, as it is a recently developed promising area of volatility model building.

6.1 Intra-Day Realized Volatility Models

The modeling of realized volatility is based on the idea of using higher frequency data to generate more accurate volatility estimates of lower frequency. Andersen and Bollerslev (1998a) introduced an alternative volatility measure, the "*realized volatility*". For P_t denoting the price of an asset at day t, let the difference of the log-prices,

$$y_{(m),t} = \log(P_t) - \log(P_{t-1/m}), \text{ where } t = 1/m, 2/m, \ldots, \tag{6.2}$$

denote the discretely observed series of continuously compounded returns with m observations per day. The realized volatility for a horizon of N days ahead is:

$$\tilde{s}^2_{t(N)} = N^{-1}\sum_{i=1}^{N}\sum_{j=1}^{m-1}\left(\log\left(P_{(j+1/m),t-i}\right)-\log\left(P_{(j/m),t-i}\right)\right)^2 . \quad (6.3)$$

Andersen at al. (2000b, 2001a, 2001b, 2001c) were the first studies that explored the distributional properties of the realized volatility. The main results are that i) although the distribution of asset returns is non-normal (highly skewed and kurtosed), the distribution of returns scaled by the realized standard deviation is approximately Gaussian and ii) the realized logarithmic standard deviation is also nearly Gaussian. The concept of the realized volatility is based on the "*integrated volatility*", which is central to the stochastic volatility option pricing in Hull and White (1987). Over an interval of length h, the integrated volatility is defined as:

$$y^2_{h,t} = \int_0^h s^2_{t-h+r}dr, \quad (6.4)$$

where s_t is the volatility of the instantaneous returns process, generated by the continuous time martingale, $d\log\left(P_t\right) = s_t dW_t$, ($W_t$ is the standard Wiener process). In the case of discrete time with a sample frequency of $h = 1/m$, $y^2_{(1/h),t}$ is an unbiased estimator of $y^2_{h,t}$. As noted by Ebens (1999) and Andersen and Bollerslev (1998a) for daily volatility forecasts, or $\left(h=1\right)$, the discretely sampled daily returns, for $\left(m=1\right)$, constitute a noisy estimator, but the accuracy improves as the sampling frequency is increasing, $\left(m\to\infty\right)$. However, the observed tick-by-tick asset prices are available only at discrete points in time and asset returns are characterized by the effect of non-synchronous trading. Thus, the sampling frequency should be as high as the market microstructure features do not induce bias to volatility estimator, i.e. Andersen and Bollerslev (1998a), Andersen et al. (1999), Andersen et al. (2000a), Andersen et al. (2001a), Areal and Taylor (2000), Kayahan et al. (2002) used a sampling frequency of 5-minites for heavily traded assets. The 5-minites sampling frequency were also used in the majority of the subsequent studies.

Ebens (1999), Giot and Laurent (2001), and Thomakos and Wang (2002) proposed the use of an ARFIMA model, in the form of (2.40), in order to fit the logarithmic realized variance. For more information and reference about applications and properties of the realized volatility and the use of intraday data see Andersen (2000), Andersen and Bollerslev (1998b), Andersen et al. (2001b), Barndorff-Nielsen and

Shephard (2002a, 2002b), Bollerslev and Wright (2001), Oomen (2001) and Taylor and Xu (1997).

7. Interpretation of ARCH Process

A number of studies have aimed at explaining the prominence of ARCH process in financial applications. Stock (1987, 1988) established the time deformation model, in which economic and calendar time proceed at different speed, and linked the relation between time deformation and ARCH models. Any economic variable evolves on an operational time scale, while in practice it is measured on a calendar time scale. The inappropriate use of calendar time scale leads to volatility clustering since relative to the calendar time, the variable may evolve quicker or slower. The time deformation model for a random variable y_t has the form:

$$\begin{aligned} y_t &= p_t y_{t-1} + \varepsilon_t \\ \varepsilon_t \mid I_{t-1} &\sim N\left(0, \sigma_t^2\right) \\ \sigma_t^2 &= a_0 + a_1 \varepsilon_{t-1}^2. \end{aligned} \tag{7.1}$$

According to Stock, when a long segment of operational time elapsed during a unit of calendar time, p_t is small and σ_t^2 is large. In order words, the time varying autoregressive parameter is inversely related to the conditional variance.

Mizrach (1990) developed a model in which the errors, made by the participants of the market on investing, are strongly dependent on all past errors. The highly persistence on the errors forces the volatility of asset returns to have an ARCH like structure.

Gallant et al. (1991), based on some earlier work by Clark (1973), Mandelbrot and Taylor (1967), Tauchen and Pitts (1983), and Westerfield (1977) provided a theoretical interpretation of ARCH effect. Let us assume that the asset returns are defined by a stochastic number of intra-period price revisions so that they can be decomposed to:

$$y_t = \mu_t + \sum_{i=1}^{\omega_t} \zeta_i \,, \tag{7.2}$$

where μ_t is the forecastable component, $\zeta_i \overset{i.i.d.}{\sim} N\left(0, s^2\right)$ denotes the incremental changes and ω_t is the number of times new information comes to the market in time t.

ω_t is an unobservable random variable and is independent of the incremental changes. In such a case, the asset returns are not normally distributed, as their distribution is a mixture of normal distributions. Rewriting the equation (7.2) as: $y_t = \mu_t + s^2\sqrt{\omega_t}\,z_t$, with z_t, $t = 1,2,...$ as i.i.d. standard normal variables, the y_t conditional on any information set, ω_t and I_{t-1}, is normally distributed:

$$y_t \,|\, (\omega_t, I_{t-1}) \sim N(\mu_t, s^2\omega_t). \tag{7.3}$$

However, the knowledge of information that flows into the market is an unrealistic assumption. Hence, the y_t conditional on the information set available to the market participants is:

$$y_t \,|\, I_{t-1} \sim N(\mu_t, s^2 E_{t-1}(\omega_t)). \tag{7.4}$$

Note that the conditional kurtosis, $3E_{t-1}(\omega_t^2)/E_{t-1}(\omega_t)^2$, exceeds 3, as in the ARCH process where the innovation, ε_t, always has fatter tails than its unconditional normal distribution:

$$E(\varepsilon_t^4)/E(\varepsilon_t^2)^2 \geq 3. \tag{7.5}$$

Lamoureux and Lastrapes (1990) assumed that the number of information arrivals is serially correlated and used the daily trading volume as a proxy variable for the daily information that flows into the stock market. Hence, ω_t can be expressed as an autoregressive process:

$$\begin{gathered} \omega_t = b_0 + \sum_{i=1}^{\kappa} b_i \omega_{t-i} + z_t \\ z_t \overset{iid}{\sim} N(0,1) \end{gathered} \tag{7.6}$$

From (7.4) we know that $E((y_t - \mu_t)^2 \,|\, I_{t-1}) = s^2\omega_t$, thus (7.6) becomes

$$E((y_t - \mu_t)^2 \,|\, I_{t-1}) = s^2 b_0 + \sum_{i=1}^{\kappa} b_i E((y_{t-i} - \mu_{t-i})^2 \,|\, I_{t-i-1}) + s^2 z_t. \tag{7.7}$$

The structure in (7.7) expresses the persistence in conditional variance, a characteristic that is captured by the ARCH process. Lamoureux and Lastrapes (1990) used the trading volume as a proxy variable for ω_t. Including the daily trading volume, V_t, as an exogenous variable in the GARCH(1,1) model, they found that its coefficient was highly significant whereas the ARCH coefficients became negligible:

$$\sigma_t^2 = a_0 + a_1\varepsilon_t^2 + b_1\sigma_t^2 + \delta V_t \,. \tag{7.8}$$

The heteroscedastic mixture model assumes that $\delta > 0$ and that the persistence of variance as measured by $a_1 + b_1$ should become negligible. Their work provided empirical evidence that the ARCH process is a manifestation of the time dependence on the rate of information arrival to the market.

Brailsford (1996) and Pyun et al. (2000) applied versions of the heteroscedastic mixture model and reported that the degree of persistence reduced as a proxy for information arrival enters into the variance equation. On the other hand, a number of studies (i.e. Abhyankar (1995), Bessembinder and Seguin (1993), Najand and Yung (1991), Locke and Sayers (1993), Sharma et al. (1996)) tested the mixture of distributions hypothesis, for various sets of data, and found that the ARCH coefficients remain statistically significant even after a trading volume is included as an exogenous variable in the model. This contradiction forced Miyakoshi (2002) to reexamine the relation between ARCH effects and rate of information arrival to the market. By using data from the Tokyo Stock Exchange, Miyakoshi showed that for periods with important market announcements, the trading volume affects the return volatility and the ARCH coefficients become negligible, while for periods which lack of "*big news*" the ARCH structure characterizes the conditional variance, adequately. The mixture of distributions hypothesis was also reexamined by Luu and Martens (2002) in the context of "*realized volatility*".

Engle et al. (1990a) evaluated the role of the information arrival process in the determination of volatility in a multivariate framework providing a test of two hypotheses: heat waves and meteor showers. Using meteorological analogies, they supposed that information follows a process like a heat wave so that a hot day in New York is likely to be followed by another hot day in New York but not typically by a hot day in Tokyo. On the other hand, a meteor shower in New York, which rains down on the earth as it turns, will almost surely be followed by one in Tokyo. Thus, the heat wave hypothesis is that the volatility has only country specific autocorrelation, while the meteor shower hypothesis states that volatility in one market spills over to the next. They examined intra daily volatility in the foreign exchange markets, focusing on time periods corresponding to the business hours of different countries. Their research based on the Yen/Dollar exchange rate while the Tokyo, European and New York market are open. They found that the foreign news was more important than the past domestic news. So,

the major effect is more like a meteor shower, i.e. Japanese news had a greater impact on the volatility of all markets except the Tokyo market. This is interpreted as evidence that volatility in part arises from trading rather than purely from news. Conrad et al. (1991), Pyun et al. (2000) and Ross (1989) examined the volatility spillover effect across large and small capitalization companies. The main finding is that volatility propagates asymmetrically in sense that the effect of shocks of larger firms on the volatility of smaller companies is more significant than that from smaller firms to larger companies.

Bollerslev and Domowitz (1991) showed how the actual market mechanisms may themselves result in a very different temporal dependence in the volatility of transaction prices, with a particular automated trade execution system inducing a very high degree of persistence in the conditional variance process.

Alternative expositions for theoretical evidence on the sources of ARCH effect have been presented by Attanasio and Wadhwani (1989), Backus et al. (1989), Brock and Kleidon (1990), Diebold and Pauly (1988), Domowitz and Hakkio (1985), Engle and Susmel (1990), Giovannini and Jorion (1989), Hodrick (1989), Hong and Lee (2001), Hsieh (1988), Lai and Pauly (1988), Laux and Ng (1993), Ng (1988), Schwert (1989a), Smith (1987) and Thum (1988). Nelson (1990b) was the first to show how ARCH models can emerge from diffusion processes. The problem of estimation of discretely sampled diffusions, such as ARCH processes, and their relationship with continuous time models has also been considered in the literature (see, e.g., Aït-Sahalia (2001, 2002), and references therein).

References

Abhyankar, A. (1995). Trading-round-the-clock: Return, Volatility, and Volume Spillovers in the Eurodollar Futures Market. *Pacific-Basin Finance Journal,* 3, 75-92

Aït-Sahalia, Y. (2001). Telling from Discrete Data Whether the Underlying Continuous Time Model is a Diffusion. *NBER Working Paper*, 8504, October 2001

Aït-Sahalia, Y. (2002). Maximum Likelihood Estimation of Discretely Sampled Diffusions: A Closed-Form Approximation Approach. *Econometrica*, 70(1), 223-262

Alexander, C. (2000). A Primer on the Orthogonal GARCH Model. *University of Reading, Working Paper*, February 2000

Andersen, T.G. (2000). Some Reflections on Analysis of High-Frequency Data. *Journal of Business and Economic Statistics,* 18(2), 146-153

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

Andersen, T.G., and T. Bollerslev (1998b). DM-Dollar Volatility: Intraday Activity Patterns, Macroeconomic Announcements, and Longer-Run Dependencies*. Journal of Finance*, 53, 219-265

Andersen, T.G., T. Bollerslev and J. Cai (2000a). Intraday and Interday Volatility in the Japanese Stock Market. *Journal of International Financial Markets, Institutions and Money*, 10, 107-130

Andersen, T.G., T. Bollerslev and F.X. Diebold (2002). Parametric and Nonparametric Volatility Measurement. *Handbook of Financial Econometrics*, (eds.) Yacine Aït-Sahalia and Lars Peter Hansen, Amsterdam, North Holland, forthcoming

Andersen, T.G., T. Bollerslev, F.X. Diebold, and H. Ebens (2001c). The Distribution of Stock Return Volatility. *Journal of Financial* Economics, 61, 43-76

Andersen, T.G., T. Bollerslev, F.X. Diebold, and P. Labys (2000b). Exchange Rate Returns Standardized by Realized Volatility are (Nearly) Gaussian. *Multinational Finance Journal*, 4, 159-179

Andersen, T. G., T. Bollerslev, F.X. Diebold, and P. Labys (2001a). The Distribution of Exchange Rate Volatility. *Journal of the American Statistical Association*, 96, 42-55

Andersen, T.G., T. Bollerslev, F.X. Diebold, and P. Labys (2001b). Modeling and Forecasting Realized Volatility. *NBER Working Paper*, No. 8160

Andersen, T.G., T. Bollerslev and S. Lange (1999). Forecasting Financial Market Volatility: Sample Frequency vis-à-vis Forecast Horizon. *Journal of Empirical Finance*, 6, 457-477

Areal N. and S. Taylor (2000). The Realized Volatility of FTSE-100 Futures Prices. *Department of Accounting and Finance, Lancaster University*, Manuscript

Atchison, M., K. Butler and R. Simonds (1987). Non-synchronous Security Trading and Market Index Autocorrelation. *Journal of Finance*, 42, 111-118

Attanasio, O.P. and S. Wadhwani (1989). Risk and the Predictability of Stock Market Returns. *Stanford University*, Manuscript

Baba, Y., R.F. Engle, D.F. Kraft and K.F. Kroner (1990). Multivariate Simultaneous Generalized ARCH. *Department of Economics, University of California*, San Diego, Mimeo

Bachellier, L. (1900). Theory of Speculation. In: P. Cootner (ed.), The Random Character of Stock Market Prices (1964), *MIT Press*, Cambridge

Backus, D.K., A.W. Gregory and S.E. Zin (1989). Risk Premiums in the Term Structure: Evidence from Artificial Economics*. Journal of Monetary Economics*, 24, 371-399

Bai, J. and S. Ng (2001). A Consistent Test for Conditional Symmetry in Time Series Models. *Journal of Econometrics*, 103, 225-258

Baillie, R.T. and T. Bollerslev (1989). The Message in Daily Exchange Rates: A Conditional Variance Tale*. Journal of Business and Economic Statistics*, 7, 297-305

Baillie, R.T. and T. Bollerslev (1990). A Multivariate Generalized ARCH Approach to Modeling Risk-Premia in Forward Foreign Exchange Rate Markets. *Journal of International Money and Finance*, 16, 109-124

Baillie, R.T., T. Bollerslev and H.O. Mikkelsen (1996). Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 74, 3-30

Baillie, R.T. and R.J. Myers (1991). Bivariate GARCH Estimation of Optimal Commodity Futures Hedge. *Journal of Applied Econometrics*, 16, 109-124

Barndorff-Nielsen, O.E., E. Nicolato and N. Shephard (2001). Some Recent Developments in Stochastic Volatility Modelling. forthcoming

Barndorff-Nielsen, O.E., and N. Shephard (2002a). How Accurate Is the Asymptotic Approximation to the Distribution of Realised Variance? *Identification and Inference for Econometric Modes,* eds. D.W.F. Andrews, J.L. Powell, P.A. Ruud and J.H. Stock, Cambridge University Press, forthcoming

Barndorff-Nielsen, O.E., and N. Shephard (2002b). Econometric Analysis of Realised Volatility and its Use in Estimating Stochastic Volatility Models. *Journal of the Royal Statistical Society, Series B,* 64, 253-280

Bates, C. and H. White (1988) Efficient Instrumental Variables Estimation of Systems of Implicit Heterogeneous Nonlinear Dynamic Equations with Non-spherical Erros. In: W.A. Barnett, E.R. Berndt, and H. White (eds.), *Dynamic Econometric Modeling, Cambridge University Press*, Cambridge

Beckers, S. (1983). Variances of Security Price Returns Based on High, Low, and Closing Prices. *Journal of Business*, 56, 97-109

Beine, M., A. Bénassy-Quéré and C. Lecourt (2002). Central Bank Intervention and Foreign Exchange Rates: New Evidence from FIGARCH Estimations. *Journal of International Money and Finance*, 21, 115-144

Bekaert, G. and G. Wu (1997). Asymmetric Volatility and Risk in Equity Markets. *NBER*, 6022, April 1997

Bera, A.K., P. Garcia and J.S. Roh (1991). Estimation of Time Varying Hedge Ratios for Agricultural Commodities: BGARCH and Random Coefficient Approaches. *University of Illinois at Urbana Champaign, Department of Economics*, Mimeo

Bera, A.K. and M.L. Higgins (1993). ARCH Models: Properties, Estimation and Testing. *Journal of Economic Surveys*, 7, 305-366

Bera, A.K. and S. Kim (1996). Testing Constancy of Correlation with an Application to International Equity Returns. *Mimeo, Center for International Business Education and Research (CIBER),* working paper 96-107, University of Illinois, Urbana-Champaign

Bera, A.K. and S. Lee (1990). On the Formulation of a General Structure for Conditional Heteroskedasticity. *University of Illinois*, Mimeo

Berkes, I. and L. Horváth (2003). The Rate of Consistency of the Quasi-Maximum Likelihood Estimator. *Statistics and Probability Letters,* 61, 133-143

Berkes, I., L. Horváth and P. Kokoszka (2003). GARCH Processes: Structure and Estimation. *Bernouli*, forthcoming

Berndt, E., B. Hall, R. Hall, and J. Hausman (1974). Estimation and Inference in Nonlinear Structural Models. *Annals of Economic and Social Measurement*, 3, 653-665

Bessembinder, H. and P. Seguin (1993). Price Volatility, Trading Volume, and Market Depth: Evidence From the Futures Market. *Journal of Financial and Quantitative Analysis*, 28, 21-39

Black, F. (1976). Studies of Stock Market Volatility Changes. Proceedings of the American Statistical Association*, Business and Economic Statistics Section*, 177-181

Blattberg, R. and N. Gonedes (1974). A Comparison of Stable and Student's t Distributions as Statistical Models for Stock Prices. *Journal of Business*, 47, 244-280

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31, 307–327

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

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

Bollerslev, T., R. Chou and K.F. Kroner (1992). ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence. *Journal of Econometrics*, 52, 5-59

Bollerslev, T., I. Domowitz (1991). Price Volatility Spread Variability and the Role of Alternative Market Mechanisms. *Review of Futures Markets*

Bollerslev, T., R.F. Engle and D. Nelson (1994). ARCH Models, in *Handbook of Econometrics*, Volume 4, eds. R. Engle and D. McFadden, Elsevier Science, Amsterdam, 2959-3038

Bollerslev, T., R.F. Engle and J.M. Wooldridge (1988). A Capital Asset Pricing Model with Time-Varying Covariances. *Journal of Political Economy*, 96, 116-131

Bollerslev, T., H.O. Mikkelsen (1996). Modeling and Pricing Long-Memory in Stock Market Volatility. *Journal of Econometrics*, 73, 151-184

Bollerslev, T. and J.M. Wooldridge (1992). Quasi-maximum Likelihood Estimation and Inference in Dynamic Models with Time-Varying Covariances. *Econometric Reviews*, 11, 143-172

Bollerslev, T. and J.H. Wright (2001). Volatility Forecasting, High-Frequency Data, and Frequency Domain Inference. *Review of Economics and Statistics*, 83, 596-602

Bougerol, P. and N. Picard (1992). Stationarity of GARCH processes and of Some Non-negative Time Series. *Journal of Econometrics*, 52, 115-128

Box, G.E.P. and D.R. Cox (1964). An Analysis of Transformations. *Journal of the Royal Statistical Society*, B 26, 211-243

Box, G.E.P. and G.C. Tiao (1973). Bayesian Inference in Statistical Analysis. *Reading, Mass*

Brailsford, T.J. (1996). The Empirical Relationship Between Trading Volume, Returns, and Volatility. *Accounting and Finance*, 35, 89-111

Braun, P.A., D.B. Nelson and A.M. Sunier (1995). Good News, Bad News, Volatility, and Betas. *Journal of Finance*, 50,1575-1603

Brock, W.A., W.D. Dechert and J.A. Scheinkman (1987). A Test for Independence Based on the Correlation Dimension. *Department of Economics, University of Wisconsin*, Madison, WI. SSRI, Working Paper no. 8702

Brock, W.A., W.D. Dechert, J.A. Scheinkman and B. LeBaron (1996). A Test for Independence Based on the Correlation Dimension. *Econometric Reviews*, 15, 3, 197-235

Brock, W.A., D.A. Hsieh and B. LeBaron (1991). Nonlinear Dynamics, Chaos, and Instability: Statistical Theory and Economic Evidence. *MIT Press*, Cambridge

Brock, W.A., and A.W. Kleidon (1990). Exogenous Demand Stocks and Trading Volume: A Model of Intraday Bids and Asks. *University of Wisconsin*, Manuscript

Brockwell, P.J. and R.A. Davis (1991). Time Series: Theory and Methods. *Springer, New York*

Brooks, C., S.P. Burke and G. Persand (2001). Benchmarks and the Accuracy of GARCH Model Estimation. *International Journal of Forecasting*, 17, 45-56

Brooks, C., R.W. Faff, M.D. McKenzie and H. Mitchell (2000). A Multi-country Study of Power ARCH Models and National Stock Market Returns. *Journal of International Money and Finance*, 19, 377-397

Brorsen, B.W. and S.R. Yang (1994). Nonlinear Dynamics and the Distribution of Daily Stock Index Returns. *Journal of Financial Research,* 17(2), 187-203

Brown, B.W. and A.V. Ligeralde (1990). Conditional Heteroskedasticity in Overlapping Prediction Models. *Rice University*, Manuscript

Brunetti, C. and C.L. Gilbert (1998). A Bivariate FIGARCH Model of Crude Oil Price Volatility. *Paper 390, Department of Economics, QMW, London*

Brunetti, C. and C.L. Gilbert (2000). Bivariate FIGARCH and Fractional Cointegration. *Journal of Empirical Finance*, 7, 509-530

Bühlmann P. and A.J. McNeil (2002) An Algorithm for Non-parametric GARCH Modelling. *Computational Statistics and Data Analysis*, 40, 665-683

Cai, J. (1994). A Markov Model of Switching-Regime ARCH. *Journal of Business and Economic Statistics*, 12, 309-316

Calzorari, G. and G. Fiorentini (1998). A Tobit Model with GARCH Errors. *Econometric Reviews*, 17, 85-104

Campbell, J., A. Lo, and A.C. MacKinlay (1997). The Econometrics of Financial Markets. *New Jersey. Princeton University Press*

Cecchetti, S.G., R.E. Cumby and S. Figlewski (1988). Estimation of the Optimal Futures Hedge. *Review of Economics and Statistics*, 70, 623-630

Chan, K.C. (1988). On the Contrarian Investment Strategy. *Journal of Business*, 61, 147-163

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

Chopra, N., J. Lakonishok and J.R. Ritter (1992). Measuring Abnormal Performance: Do Stocks Overreact? *Journal of Financial Economics*, 31, 235-268

Choudhry, T. (1995). Integrated-GARCH and Non-stationary Variances: Evidence from European Stock Markets During the 1920s and 1930s. *Economics Letters*, 48, 55-59

Christie, A.A. (1982). The Stochastic Behavior of Common Stock Variances: Value, Leverage and Interest Rate Effects. *Journal of Financial Economics*, 10, 407-432

Christodoulakis, G.A. and S.E. Satchell (2002). Correlated ARCH (CorrARCH): Modelling the Time-Varying Conditional Correlation Between Financial Asset Returns. *European Journal of Operational Research,* 139, 351-370

Clark, P.K. (1973). A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices. *Econometrica*, 41, 135-156

Cohen, K., G. Hawawini, S. Maier, R. Schwartz, and D. Whitcomb (1979). On the Existence of Serial Correlation in an Efficient Securities Market. *TIMS Studies in the Management Sciences*, 11, 151-168

Cohen, K., G. Hawawini, S. Maier, R. Schwartz, and D. Whitcomb (1983). Friction in the Trading Process and the Estimation of Systematic Risk. *Journal of Financial Economics*, 12, 263-278

Cohen, K., S. Maier, R. Schwartz and D. Whitcomb (1978). The Returns Generation Process, Returns Variance, and the Effect of Thinness in Securities Markets. *Journal of Finance*, 33, 149-167

Conrad, J., M.N. Gultekin and G. Kaul (1991). Asymmetric Predictability of Conditional Variances. *Review of Financial Studies*, 4, 597-662

Conrad, J. and G. Kaul (1993). Long-term Market Overreaction or Biases in Computed Returns. *Journal of Finance*, 48, 39-63

Dacarogna, M.M., U.A. Muller, R.J. Nagler, R.B. Olsen and O.V. Pictet (1993). A Geographical Model for the Daily and Weekly Seasonal Volatility in the Foreign Exchange Market. *Journal of International Money and Finance*, 12, 413-438

Davis, P.J. (1965). The Gamma Function and Related Functions. In M. Abramowitz and N. Stegun (eds.), In Chapter 6 of *Handbook of Mathematical Functions* (New York), 253-294

Day, T.E. and C.M. Lewis (1988). The Behaviour of the Volatility Implicit in the Prices of Stock Index Options. *Journal of Financial Economics*, 22, 103-122

Day, T.E. and C.M. Lewis (1992). Stock Market Volatility and the Information Content of Stock Index Options. *Journal of Econometrics*, 52, 267-287

DeBondt, W.F.M. and R.H. Thaler (1985). Does the Stock Market Overreact? *Journal of Finance,* 40, 793-805

DeBondt, W.F.M. and R.H. Thaler (1987). Further Evidence on Investor Overreaction and Stock Market Seasonality. *Journal of Finance,* 42, 557-581

DeBondt, W.F.M. and R.H. Thaler (1989). A Mean-Reverting Walk Down Wall Street. *Journal of Economic Perspectives,* 3, 189-202

Dempster, A.P., N.M. Laird and D.R. Rubin (1977). Maximum Likelihood From Incomplete Data Via the EM Algorithm. *Journal of the Royal Statistical Society, Series B,* 39, 1-22

Devaney, M. (2001). Time Varying Risk Premia For Real Estate Investent Trusts: A GARCH-M Model. *Quartely Review of Economics and Finance*, 41, 335-346

Diebold, F.X. and M. Nerlove (1989). The Dynamics of Exchange Rate Volatility: A Multivariate Latent Factor ARCH Model. *Journal of Applied Econometrics*, 4, 1-21

Diebold, F.X. and P. Pauly (1988). Endogenous Risk in a Portfolio Balance Rational Expectations Model of the Deutschmark-Dollar Rate. *European Economic Review*, 32, 27-53

Dimson, E., (1979). Risk Measurement When Shares Are Subject to Infrequent Trading. *Journal of Financial Economics*, 7, 197-226

Ding, Z. and R.F. Engle (2001) Large Scale Conditional Covariance Matrix Modeling, Estimation and Testing, *Working Paper, Department of Finance,* Stern School of Business, New York University, May 2001

Ding, Z., C.W.J. Granger, and R.F. Engle (1993). A Long Memory Property of Stock Market Returns and a New Model. *Journal of Empirical Finance*, 1, 83-106

Domowitz, I. and C.S. Hakkio (1985). Conditional Variance and the Risk Premium in the Foreign Exchange Market. *Journal of International Economics*, 19, 47-66

Drost, F.C., T.E. Nijman and B.J.M. Werker (1998). Estimation and Testing in Models Containing Both Jumps and Conditional Heteroskedasticity. *Journal of Business and Economic Statistics,* 16(2), 237-243

Duan, J.C. (1996). A Unified Theory of Option Pricing Under Stochastic Volatility – From GARCH to Diffusion. *Hong Kong University of Science and Technology*, Working Paper

Dudley, R.M. (1989). Real Analysis and Probability, *Pacific Grove, California*

Dueker, M.J. (1997). Markov Switching in GARCH Processes and Mean-Reverting Stock Market Volatility. *Journal of Business and Economic Statistics*, 15, 26-34

Ebens, H. (1999). Realized Stock Volatility. *Department of Economics, Johns Hopkins University,* Working Paper, 420

Einstein, A. (1905). Ueber die von der molecular-kinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annalen der Physik*, 17, 549-560

Elyasiani, E. and I. Mansur (1998). Sensitivity of Bank Stock Returns Distribution to Changes in the Level and Volatility of Interest Rates: A GARCH-M Model. *Journal of Banking and Finance*, 22, 535-563.

Engle, R.F. (1982). Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation. *Econometrica*, 50, 987–1008

Engle, R.F. (1983). Estimates of the Variance of US Inflation Based on the ARCH Model. *Journal of Money Credit and Banking*, 15, 286–301

Engle, R.F. (1987). Multivariate ARCH with Factor Structures – Cointegration in Variance. *University of California, San Diego*, Discussion Paper, 87-27

Engle, R.F. (1990). Discussion: Stock Market Volatility and the Crash of '87. *Review of Financial Studies*, 3, 103-106

Engle, R.F. (2000). Dynamic Conditional Correlation - A Simple Class of Multivariate GARCH Models. *Universiy of California*, 2000(09), May, 2000

Engle, R.F. and T. Bollerslev (1986). Modelling the Persistence of Conditional Variances. *Econometric Reviews*, 5(1), 1-50

Engle, R.F. and G. González-Rivera (1991). Semiparametric ARCH Models. *Journal of Business and Economic Statistics*, 9, 345-359

Engle, R.F., C.W.J. Granger and D. Kraft (1986). Combining Competing Forecasts of Inflation Using a Bivariate ARCH Model. *Journal of Economic Dynamics and Control*, 8, 151-165

Engle, R.F., T. Ito and W.L. Lin (1990a). Meteor Showers or Heat Waves? Heteroskedastic Intra-Daily Volatility in the Foreign Exchange Market. *Econometrica*, 58, 525-542

Engle, R.F. and D. Kraft (1983). Multiperiod Forecast Error Variances of Inflation Estimated from ARCH Models, in A. Zellner (ed.), *Applied Time Series Analysis of Economic Data*, (Washington, DC), 293–302

Engle, R.F. and K.F. Kroner (1995). Multivariate Simultaneous Generalized ARCH. *Econometric Theory*, 11, 122-150

Engle, R.F. and G.G.J. Lee (1993). A Permanent and Transitory Component Model of Stock Return Volatility. *University of California, San Diego, Department of Economics*, Discussion Paper, 9244

Engle, R.F., D.M. Lilien and R.P. Robins (1987). Estimating Time Varying Risk Premia in the Term Structure: The ARCH-M Model. *Econometrica*, 55, 391–407

Engle, R.F. and V.K. Ng (1993). Measuring and Testing the Impact of News on Volatility. *Journal of Finance*, 48, 1749-1778

Engle, R.F., V.K. Ng and M. Rothschild (1990b). Asset Pricing with a Factor ARCH Covariance Structure: Empirical Estimates for Treasury Bills. *Journal of Econometrics*, 45, 213-238

Engle, R.F. and K. Sheppard (2001). Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH. *University of California*, 2001(15), September, 2001

Engle, R.F., and R. Susmel (1990). Intraday Mean and Volatility Relations Between US and UK Stock Market Returns. *University of California, San Diego*, Manuscript

Fama, E.F. (1965). The Behaviour of Stock Market Prices. *Journal of Business*, 38, 34-105

Fama, E.F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal of Finance*, 25, 383-417

Fernandez, C. and M. Steel (1998). On Bayesian Modeling of Fat Tails and Skewness. *Journal of the American Statistical Association*, 93, 359-371

Ferson, W.E. (1989). Changes in Expected Security Returns, Risk and the Level of Interest Rates. *Journal of Finance*, 44, 1191-1218

Fisher, L. (1966). Some New Stock Market Indices. *Journal of Business*, 39, 191-225

Fornari, F. and A. Mele (1995). Sign- and Volatility-Switching ARCH Models: Theory and Applications to International Stock Markets. *Banca d'Italia, Discussion Paper*, 251

Fornari, F. and A. Mele (1996). Modeling the Changing Asymmetry of Conditional Variances. *Economics Letters*, 50, 197-203

Fornari, F., C. Monticelli, M. Pericoli and M. Tivegna (2002). The Impact of News on the Exchange Rate of the Lira and Long-term Interest Rates. *Economic Modelling*, 19, 611-639

French, K.R. and R. Roll (1986). Stock Return Variances: The Arrival of Information and the Reaction of Traders. *Journal of Financial Economics*, 17, 5-26

French, K.R., G.W. Schwert and R.F. Stambaugh (1987). Expected Stock Returns and Volatility. *Journal of Financial Economics*, 19, 3-29

Friedmann, R. and W.G. Sanddorf-Köhle (2002). Volatility Clustering and Non-trading Days in Chinese Stock Markets. *Journal of Economics and Business*, 54, 193-217

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

Gallant, A.R., P.E. Rossi, and G. Tauchen (1992). Stock Prices and Volume. *Review of Financial Studies*, 5, 199-242

Gallant, A.R., P.E. Rossi, and G. Tauchen (1993). Nonlinear Dynamic Structures. *Econometrica*, 61, 871-907

Gallant, A.R. and G. Tauchen (1989). Semi Non-Parametric Estimation of Conditional Constrained Heterogeneous Processes: Asset Pricing Applications. *Econometrica*, 57, 1091-1120

Garman, M.B. and M.J. Klass (1980). On the Estimation of Security Price Volatilities From Historical Data. *Journal of Business*, 53(1), 67-78

Geweke, J. (1986). Modeling the Persistence of Conditional Variances: A Comment. *Econometric Reviews*, 5, 57-61

Geweke, J. (1988a). Comments on Poirier: Operational Bayesian Methods in Econometrics. *Journal of Economic Perspectives*, 2, 159-166

Geweke, J. (1988b). Exact Inference in Models with Autoregressive Conditional Heteroskedasticity, in: W.A. Barnett, E.R. Berndt, and H. White (eds.*), Dynamic Econometric Modeling*, Cambridge University Press, Cambridge

Geweke, J. (1989). Exact Predictive Densities in Linear Models with Arch Distrubances. *Journal of Econometrics*, 44, 307-325

Ghysels, E., A. Harvey and E. Renault (1996). Stochastic Volatility, in G.S. Maddala (ed.), *Handbook of Statistics, Vol. 14, Statistical Methods in Finance*, 119-191, Amsterdam, North Holland

Giot, P. and S. Laurent (2001). Modelling daily Value-at-Risk using realized volatility and ARCH type models, *METEOR Research Memorandum*, April, RM/01

Giovannini, A. and P. Jorion (1989). The Time Variation of Risk and Return in the Foreign Exchange and Stock Markets. *Journal of Finance*, 44, 307-325

Giraitis, L. and P.M. Robinson (2000). Whittle Estimation of ARCH Models. *Econometric Theory*, 17,608-631

Glosten, L., R. Jagannathan, and D. Runkle (1993). On the Relation Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *Journal of Finance*, 48, 1779–1801

González-Rivera, G. (1996). Smooth Transition GARCH Models. *University of California, Riverside, Department of Economics,* Working Paper

Gourieroux, C. (1997). ARCH models and Financial Applications. *Springer-Verlag, New York*

Gouriéroux, C. and A. Monfort (1992). Qualitative Threshold ARCH Models. *Journal of Econometrics*, 52, 159-199

Granger, C.W.J. (1980). Long Memory Relationships and the Aggregation of Dynamic Models. *Journal of Econometrics*, 14, 227-238

Granger, C.W.J. (1981). Some Properties of Time Series Data and their Use in Econometric Model Specification. *Journal of Econometrics*, 16, 121-130

Granger, C.W.J. and R. Joyeux (1980). An Introduction to Long Memory Time Series Models and Fractional Differencing. *Journal of Time Series Analysis*, 1, 15-39

Hagerman, R. (1978). More Evidence on the Distribution of Security Returns. *Journal of Finance*, 33, 1213-1220

Hagerud, G. E. (1996). A Smooth Transition Arch Model for Asset Returns. *Department of Finance, Stockholm School of Economics*, Working Paper Series in Economics and Finance, 162

Hall, P. and Q. Yao (2003). Inference in ARCH and GARCH Models with Heavy-Tailed Errors. *Econometrica*, 71, 285-317

Hamilton, J. D. (1989). A new Approach to the Economic Analysis of Non-stationary Time Series and the Business Cycle. *Econometrica*, 57, 357-384

Hamilton, J.D. and R. Susmel (1994). Autoregressive Conditional Heteroskedasticity and Changes in Regime. *Journal of Econometrics*, 64, 307-333

Hansen, B.E. (1994). Autoregressive Conditional Density Estimation. *International Economic Review*, 35, 705-730

Harvey, A.C. (1981). The Econometric Analysis of Time Series*, Oxford*

Harvey, A.C. and N. Shephard (1993). The Econometrics of Stochastic Volatility. *London School of Economics,* Discussion Paper, 166.

Harvey, A.C., E. Ruiz and E. Sentana (1992). Unobserved Component Time Series Models with ARCH Disturbances. *Journal of Econometrics*, 52, 129-157

Hentschel, L. (1995). All in the Family: Nesting Symmetric and Asymmetric GARCH Models. *Journal of Financial Economics*, 39, 71-104

Higgins, M.L. and A.K. Bera (1992). A Class of Nonlinear ARCH models. *International Economic Review*, 33, 137-158

Hodrick R.J. (1989). Risk, Uncertainty, and Exchange Rates. *Journal of Monetary Economics*, 23, 433-459

Hong Y., and J. Lee (2001). One-Sided Testing for ARCH Effects Using Wavelets. *Econometric Theory*, 17(6), 1051-1081

Hosking, J.R.M. (1981). Fractional Differencing. *Biometrica*, 68, 165-176

Hsieh, D.A. (1989). Modeling Heteroscedasticity in Daily Foreign-Exchange Rates. *Journal of Business and Economic Statistics*, 7, 307-317

Hull, J., and A. White (1987). The Pricing of Options on Assets with Stochastic Volatility, *Journal of Finance,* 42, 281-300

Hurst, H. (1951). The Long Term Storage Capacity of Reservoirs. *Transactions of the American Society of Civil Engineers*, 116, 770-799

Hwang, Y. (2001). Asymmetric Long Memory GARCH in Exchange Return. *Economics Letters*, 73, 1-5

Jacquier, E., Polson, N. and P. Rossi (1994). Bayesian Analysis of Stochastic Volatility Models. *Journal of Business and Economic Statistics*, 12, 4, 371-417

Jacquier, E., Polson, N. and P. Rossi (1999). Stochastic Volatility: Univariate and Mulrivariate Extensions. *CIRANO, Scientific Series*, 99s, 26

Jondeau, E. and M. Rockinger (2001). Gram-Charlier Densities. *Journal of Economic Dynamics & Control*, 25, 1457-1483

Jorion, P. (1988). On Jump Processes in the Foreign Exchange and Stock Markets. *Review of Financial Studies*, 1, 427-445

Kayahan B., T. Saltoglu and T. Stengos (2002). Intra-Day Features of Realized Volatility: Evidence from an Emerging Market. *International Journal of Business and Economics*, 1(1), 17-24

Kendal, M.G. and A. Stuart (1969). The advanced Theory of Statistics, vol. 1, 3rd eds, *Charles Griffin, London*

Kim, C.M. (1989). Volatility Effect on Time Series Behaviour of Exchange Rate Changes. *Korea Institute for International Economic Policy*, Working Paper

Kim, J. (2000). The Relationship Between The Monetary Regime and Output Volatility: A Multivariate GARCH-M Model of the Japanese Experience, 1919-1996. *Japan and the World Economy*, 12, 49-69

King, M., E. Sentana and S. Wadhwani (1994). Volatility and Links Between National Stock Markets. *Econometrica*, forthcoming

Kodres, L.E. (1988). Tests of Unbiasedness in Foreign Exchange Futures Markets: The Effects of Price Limits. *Review of Futures Markets*, 7, 138-166

Kodres, L.E. (1993). Tests of Unbiasedness in the Foreign Exchange Futures Markets: An Examination of Price Limits and Conditional Heteroskedasticity. *Journal of Business*, 66(3), 463-490

Kroner, K.F. and S. Claessens (1991). Optimal Currency Composition of External Debt: Applications to Indonesia and Turkey. *Journal of International Money and Finance*, 10, 131-148

Kroner, K.F. and W.D. Lastrapes (1991). The Impact of Exchange Rate Volatility on International Trade: Estimates Using the GARCH-M Model. *University of Arizona*, Manuscript

Kroner, K.F. and V.K. Ng (1998). Modeling Asymmetric Comovements of Asset Returns. *Review of Financial Studies*, 11, 4, 817-844

Kroner, K.F. and J. Sultan (1991). Exchange Rate Volatility and Time Varying Hadge Ratios, in S.G. Rhee and R.P. Change (eds.), *Pacific-Basin Capital Markets Research*, Vol. II, North Holland, Amsterdam

Kroner, K.F. and J. Sultan (1993). Time Varying Distribution and Dynamic Hedging with Foreign Currency Futures. *Journal of Financial and Quantitative Analysis*, 28, 535-551

Kunitomo, N. (1992). Improving the Parkinson Method of Estimating Security Price Volatilities. *Journal of Business*, 65, 295-302

Lai, K.S. and P. Pauly (1988). Time Series Properties of Foreign Exchange Rates Re-examined. *University of Pennsylvania*, Manuscript

Lambert, P. and S. Laurent (2000). Modeling Skewness Dynamics in Series of Financial Data. Discussion Paper*, Institut de Statistique, Louvain-la-Neuve*

Lambert, P. and S. Laurent (2001). Modeling Financial Time Series Using GARCH-Type Models and a Skewed Student Density.*Universite de Liege,* Mimeo

Lamoureux, G.C. and W.D. Lastrapes (1990). Heteroskedasticity in Stock Return Data: Volume Versus GARCH Effects. *Journal of Finance*, 45, 221-229

Latane H.A. and R.J. Rendleman (1976). Standard Deviations of Stock Price Ratios Implied in Option Prices. *Journal of Finance*, 31, 369-381

Laux, P.A. and L.K. Ng (1993). The Sources of GARCH: Empirical Evidence From an Intraday Returns Model Incorporating Systematic and Unique risks. *Journal of International Money and Finance*, 12, 543-560

LeBaron, B. (1992). Some Relations Between Volatility and Serial Correlations in Stock Market Returns. *Journal of Business*, 65, 2, 199-219

Lebedev, N.N. (1972). Special Functions and their Applications. *New York*

Lee, S.W. (1991). Asymptotic Properties of the Maximum Likelihood Estimator of the GARCH-M and I-GARCH-M Models. Mimeo. *Department of Economics, University of Rochester*

Lee, L.F. (1999). Estimation of Dynamic and ARCH Tobit Models. *Journal of Econometrics*, 92, 355-390

Lee, J. (1999b). The Inflation and Output Variability Trade-off: Evidence from a GARCH Model. *Economics Letters*, 62, 63-67

Lee, S.W and Hansen, B.E. (1991). Asymptotic Properties of the Maximum Likelihood Estimator and Test of the Stability of Parameters of the GARCH and I-GARCH Models. Mimeo. *Department of Economics, University of Rochester*

Lee, S.W and Hansen, B.E. (1994). Asymptotic Theory for the GARCH(1,1) Quasi-Maximum Likelihood Estimator. *Econometric Theory*, 10, 29-52

Lee, T.K.Y. and Y.K. Tse (1991). Term Structure of Interest Rates in the Singapore Asian Dollar Market. *Journal of Applied Econometrics*, 6, 143-152

Li, W.K., S. Ling and M. McAleer (2001). A Survey of Recent Theoretical Results for Time Series Models with GARCH Errors. *The Institute of Social and Economic Research, Osaka University, Japan*, Discussion Paper, 545

Lien, D. and Y.K. Tse (1998). Hedging Time Varying Downside Risk. *Journal of Futures Markets,* 18, 705-722

Lin, B.H. and S.-K. Yeh (2000). On the Distribution and Conditional Heteroscedasticity in Taiwan Stock Prices. *Journal of Multinomial Financial Management*, 10, 367-395

Ling, S. and M. McAleer (2001). Necessary and Sufficient Moment Conditions for the GARCH(r,s) and Asymmetric Power GARCH(r,s) Models. *Econometric Theory*, forthcoming

Liu, S. and B.W. Brorsen (1995). Maximum Likelihood Estimation of a GARCH-stable Model. *Journal of Applied Econometrics,* 10, 273-285

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

Lo, A. and A.C. MacKinlay (1990a). An Econometric Analysis of Non-synchronous Trading. *Journal of Econometrics*, 45, 181-212

Lo, A. and A.C. MacKinlay (1990b). When Are Contrarian Profits Due to Stock Market Overreaction? *Review of Financial Studies*, 3, 175-205

Locke, P.R. and C.L. Sayers (1993). Intra-day Futures Prices Volatility: Information Effects and Variance Persistence. *Journal of Applied Econometrics,* 8, 15-30

Longin, F. and B. Solnik (1995). Is the Correlation in International Equity Returns Constant: 1960-1990? *Journal of International Money and Finance*, 14(1), 3-26

Lubrano, M. (1998). Smooth Transition GARCH Models: A Bayesian Perspective. *CORE*, Discussion Paper, 9866

Lumsdaine, R.L. (1991). Asymptotic Properties of the Quasi-maximum Likelihood Estimator in GARCH(1,1) and IGARCH(1,1) models, Mimeo, *Department of Economics, Princeton University*

Lumsdaine, R.L. (1996). Consistency and Asymptotic Normality of the Quasi-maximum Likelihood Estimator in IGARCH(1,1) and Covariance Stationary GARCH(1,1) models, *Econometrica*, 64:3, 575-596

Luu, J.C. and M. Martens (2002). Testing the Mixture of Distributions Hypothesis Using “Realized” Volatility. *Journal of Futures Markets*, forthcoming

Luukkonen, R., P. Saikkonen and T. Teräsvirta (1988). Testing Linearity Against Smooth Transition Autoregressive Models. *Biometrica*, 75, 491-499

Mandelbrot, B. (1963). The Variation of Certain Speculative Prices*. Journal of Business*, 36, 394-419

Mandelbrot, B. (1964). New Methods in Statistical Economics. *Journal of Political Economy*, LXXI, October, 421-440

Mandelbrot, B. (1982). The Fractal Geometry of Nature. W. H. Freeman, *New York*

Mandelbrot, B. and H. Taylor (1967). On the Distribution of Stock Price Differences. *Operational Research*, 15, 1057-1062

Mandelbrot, B. and J. Van-Ness (1968). Fractional Brownian Motion, Fractional Noises and Applications*. S.I.A.M. Review*, 10, 422-437

Mark, N. (1988). Time Varying Betas and Risk Premia in the Pricing of Forward Foreign Exchange Contracts. *Journal of Financial Economics*, 22, 335-354

Marquardt, D.W. (1963). An Algorithm for Least Squares Estimation of Nonlinear Parameters. *Journal of the Society for Industrial and Applied Mathematics*, 11, 431-441

McDonald, J.B. and W.K. Newey (1988). Partially Adaptive Estimation of Regression Models via the Generalized t Distribution. *Econometric Theory*, 4, 428-457

Mikosch, T. and D. Straumann (2002). Whittle Estimation in a Heavy-Tailed GARCH(1,1) Model. *Stochastic Processes and their Applications,* 100, 187-222

Milhøj, A. (1987). A Multiplicative Parameterization of ARCH Models. *University of Copenhagen, Department of Statistics*, Mimeo

Mills, T.C. (1996). Non-Linear Forecasting of Financial Time Series: An Overview and Some New Models, *Journal of Forecasting*, 15, 127-135

Mills, T.C. (1999). The Econometric Modeling of Financial Time Series, 2nd Edition, *Cambridge University Press*, UK

Mittnik, S. and M.S. Paolella (2001). Prediction of Financial Downside-Risk with Heavy Tailed Conditional Distributions. In: *Handbook of Heavy Tailed Distributions in Finance*, North-Holland

Mittnik, S., M.S. Paollela and S.T. Rachev (1998a). Uncoditional and Conditional Distributional Models for the Nikei's Index. *Asian-Pacific Financial Markets*, 5, 99-128

Mittnik, S., M.S. Paolella and S.T. Rachev (2002). Stationarity of Stable Power-GARCH Processes. *Jounral of Econometrics*, 106, 97-107

Mittnik, S., S.T. Rachev, T. Doganoglu and D. Chenyao (1999). Maximum Likelihood Estimation of Stable Paretian Models. *Mathematical and Computer Modelling*, 29, 275-293

Mittnik, S., S.T. Rachev and M.S. Paolella (1998b). Stable Paretian Modeling in Finance: Some Empirical and Theoretical Aspects. In *A Practical Guide to Heavy Tails*, (Edited by R. Adler, R. Feldman and M.S. Taqqu), Birkhäuser, Boston, 79-110

Miyakoshi, T. (2002). ARCH Versus Information Based Variances: Evidence From the Tokyo Stock Market. *Japan and the World Economy*, 14, 215-231

Mizrach, B. (1990). Learning and Conditional Heteroskedasticity in Asset Returns, *University of Pennsylvania, The Wharton School, Department of Finance,* Mimeo

Morgan, I. and R. Trevor (1997). Limit Moves as Censored Observations of Equilibrium Futures Price in GARCH Processes. *CMBF paper*, Macquarie University, 18, 1-42

Moschini, G. and R.J. Myers (2002). Testing for Constant Hedge Ratios in Commodity Markets: A Multivariate GARCH Approach. *Journal of Empirical Finance*, 9, 589-603

Müller, U.A., M.M. Dacorogna, R.D. Davé, R.B. Olsen, O.V. Pictet and J.E. VonWeizsäcker (1997). Volatilities of Different Time Resolutions – Analyzing the Dynamics of Market Components. *Journal of Empirical Finance*, 4, 213-239

Myers, R.J. (1991). Estimating Time-Varying Optimal Hedge Ratio on Futures Markets. *Journal of Futures Markets,* 11, 39-53

Najand, M. and K. Yung (1991). A GARCH Examination of Relationship Between Volume and Price Variability in Futures Markets. *Journal of Futures Markets*, 11, 613-621

Nam, K., C.S. Pyun and A.C. Arize (2002). Asymmetric Mean Reversion and Contrarian Profits: ANST-GARCH approach. *Journal of Empirical Finance*, 9, 563-588

Nelson, D. (1990a). Stationarity and Persistence in the Garch(1,1) Model. *Econometric Theory*, 6, 318-334

Nelson, D. (1990b). ARCH Models as Diffusion Approximations. *Journal of Econometrics*, 45, 7-38

Nelson, D. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica*, 59, 347-370

Nelson, D. and C.Q. Cao (1992). Inequality Constraints in the Univariate Garch Model*. Journal of Business and Economic Statistics*, 10, 229-235

Ng, V.K. (1988). Equilibrium Stock Return Distributions in a Simple Log-linear Factor Economy. *University of California, San Diego*, Manuscript

Ng, V.K., R.F. Engle and M. Rothschild (1992). A Multi-Dynamic Factor Model for Stock Returns*. Journal of Econometrics*, 52, 245-265

Noh, J., R.F. Engle and A. Kane (1994). Forecasting Volatility and Option Prices of the S&P500 Index. *Journal of Derivatives*, 17-30

Nowicka-Zagrajek, J. and A. Weron (2001). Dependence Structure of Stable R-GARCH Processes. *Probability and Mathematical Statistics*, 21, 371-380

Oedegaard, B.A. (1991). Empirical Tests of Changes in Autocorrelation of Stock Index Returns. *Carnegie Mellon University, Graduate School of Industrial Administration*, Mimeo

Officer, R. (1972). The Distribution of Stock Returns. *Journal of the American Statistical Association*, 67, 807-812

Oomen, R. (2001). Using High Frequency Stock Market Index Data to Calculate, Model and Forecast Realized Volatility. *Department of Economics, European University Institute,* Manuscript

Ozaki, T. (1980). Non-linear Time Series Models for Non-linear Random Vibrations. *Journal of Applied Probability*, 17, 84-93

Pagan, A.R. and Y. Hong (1991). Non-Parametric Estimation and the Risk Premium. In: W.A. Barnett, J. Powell, and G. Tauchen (eds.), Nonparametric and Semiparametric Methods in Econometrics and Statistics*, Cambridge University Press*, Cambridge

Pagan, A., and H. Sabau (1987). On the Inconsistency of the MLE in Certain Heteroskedastic Regression Models. *University of Rochester, Department of Economics*, Mimeo

Pagan, A.R. and G.W. Schwert (1990). Alternative Models for Conditional Stock Volatility. *Journal of Econometrics*, 45, 267-290

Panorska, A., S. Mittnik and S.T. Rachev (1995). Stable GARCH Models for Financial Time Series. *Applied Mathematics Letters*, 8, 33-37

Pantula, S.G. (1986). Modeling the Persistence of Conditional Variances: A Comment. *Econometric Reviews*, 5, 71-73

Park, T.H. and L.N. Switzer (1995). Bivariate GARCH Estimation of the Optimal Hedge Ratios for Stock Index Futures: A Note. *Journal of Futures Markets,* 15, 61-67

Parkinson, M. (1980). The Extreme Value Method for Estimating the Variance of the Rate of Return. *Journal of Business*, 53(1), 61-65

Poon, S.,-H. and C. Granger (2001). Forecasting Financial Market Volatility: A Review. *Department of Economics, University of California, San Diego,* manuscript

Praetz, P. (1972). The Distribution of Share Price Changes. *Journal of Business*, 45, 49-55

Pyun, C.S., S.Y. Lee and K. Nam (2000). Volatility and Information Flows in Emerging Equity Market: A Case of the Korean Stock Exchange. *International Review of Financial Analysis*, 9, 405-420

Rabemananjara, R. and J.M. Zakoian (1993). Threshold ARCH Models and Asymmetries in Volatility. *Journal of Applied Econometrics*, 8, 31-49

Rich, R.W., J. Raymond and J.S. Butler (1991). Generalized Instrumental Variables Estimation of Autoregressive Conditional Heteroskedastic Models. *Economics Letters*, 35, 179-185

Rogers, L.C.G. and S.E. Satchell (1991). Estimating Variance From High, Low and Closing Prices. *The Annals of Applied Probability*, 1, 504-512

Ross, S.A. (1989). Information and Volatility: The No-Arbitrage Martingale Approach to Timing and Resolution Irrelevancy. *Journal of Finance*, 44, 1-17

Ruiz, E. and A. Pérez. Asymmetric Long Memory GARCH: A Reply to Hwang's Model. *Economics Letters*, forthcoming

Schmalensee, R. and R.R. Trippi (1978). Common Stock Volatility Expectations Implied by Option Premia. *Journal of Finance*, 33(1), 129-147

Scholes, M., and J. Williams (1977). Estimating Betas from Non-Synchronous Data. *Journal of Financial Economics*, 5, 309-328

Schwert, G.W. (1989a). Why Does Stock Market Volatility Changes Over Time. *Journal of Finance*, 44, 1115-1153

Schwert, G.W. (1989b). Business Cycles, Financial Crisis, and Stock Volatility. *Carnegie Rochester Conferecne Series on Public Policy*, 39, 83-126

Schwert, G.W. (1990). Stock Volatility and the Crash of '87. *Review of Financial Studies*, 3, 77-102

Sentana, E. (1995). Quadratic ARCH Models. *Review of Economic Studies*, 62, 639-661

Sentana, E. and S. Wadhwani (1991). Feedback Traders and Stock Returns Autocorrelations: Evidence from a Century of Daily Data. *Economic Journal*, 58, 547-563

Sharma, J.L., M. Mougoue and R.R. Kamath (1996). Heteroskedasticity in Stock Market Indicator Return Data: Volume Versus GARCH Effects. *Applied Financial Economics*, 6, 337-342

Shephard, N. (1996). Statistical Aspects of ARCH and Stochastic Volatility Models. In *Time Series Models in Econometrics, Finance and Other* Fields, 1-67, (eds.) D.R. Cox, D.V. Hinkley and O.E. Barndorff-Nielsen, Chapman & Hall, London

Simon, D.P. (1989). Expectations and Risk in the Treasury Bill Market: An Instrumental Variables Approach. *Journal of Financial and Quantitative Analysis*, 24, 357-366

Smith, G.W. (1987). Endogenous Conditional Heteroskedasticity and Tests of Menu-Cost Pricing Theory. *Queen's University,* Manuscript

Stock, J.H. (1987). Measuring Business Cycle Time. *Journal of Political Economy*, 95, 1240-1261

Stock, J.H. (1988). Estimating Continuous Time Process Subject to Time Deformation. *Journal of the Statistical Association*, 83, 77-85

Tai, C.-S. (2001). A Multivariate GARCH in Mean Approach to Testing Uncovered Interest Parity: Evidence From Asia-Pacific Foreign Exchange Markets. *Quartely Review of Economics and Finance*, 41, 441-460

Tauchen, G. and M. Pitts (1983). The Price Variability-Volume Relationship on Speculative Markets. *Econometrica*, 51, 485-505

Taylor, S.J. (1986). Modeling Financial Time Series. *Wiley and Sons: New York*, NY

Taylor, S.J. (1994).Modelling Stochastic Volatility: A Review and Comparative Study. *Mathematical Finance*, 4, 183-204

Taylor, S. and X. Xu (1997). The Incremental Volatility Information in One Million Foreign Exchange Quotations. *Journal of Empirical Finance*, 4, 317-340

Teräsvirta, T. (1994). Specification, Estimation, and Evaluation of Smooth Transition Autoregressive Models. *Journal of American Statistical Association*, 89, 208-218

Thomakos, D.D. and T. Wang (2002). Realized Volatility in the Futures Markets. *Department of Economics, Florida International University,* Manuscript

Thum, F. (1988). Economic Foundations of ARCH Exchange Rate Process. *University of Texas*, Manuscript

Tse, Y.K. (1998). The Conditional Heteroskedasticity of the Yen-Dollar Exchange Rate. *Journal of Applied Econometrics*, 193, 49-55

Tse, Y.K. (2000). A Test for Constant Correlations in a Multivariate GARCH Model. *Journal of Econometrics*, 98, 107-127

Tsionas, E.G. (1999). Monte Carlo Inference in Econometric Models with Symmetric Stable Disturbances. *Journal of Econometrics*, 88, 2, 365-401

Tsionas, E.G. (2002). Likelihood-Based Comparison of Stable Paretian and Competing Models: Evidence from Daily Exchange Rates. *Journal of Statistical Computation and Simulation*, 72, 4, 341-353

Tsui, A.K. and Q. Yu (1999). Constand Conditional Correlation in a Bivariate GARCH Model: Evidence From the Stock Markets of China. *Mathematics and Computers in Simulation*, 48, 503-509

Veronesi, P. (1999). Stock Market Overreaction to Bad News in Good Times: A Rational Expectations Equilibrium Model. *Review of Financial Studies*, 12, 975-1007

Vilasuso, J. (2002). Forecasting Exchange Rate Volatility. *Economics Letters*, 76, 59-64

Vlaar, P.J.G. and F.C. Palm (1993). The Message in Weekly Exchange Rates in the European Monetary System: Mean Reversion, Conditional Heteroscedasticity, and Jumps. *Journal of Business and Economic Statistics,* 11(3), 351-360

Vries, C.G. de. (1991). On the Relation Between GARCH and Stable Processes. *Journal of Econometrics*, 48, 313-324

Wei, S.X. (1999). A Bayesian Approach to Dynamic Tobit Models. *Econometric Reviews*, 18, 417-439

Wei, S.X. (2002). A Censored GARCH Model of Asset Returns with Price Limits. *Journal of Empirical Finance*, 9, 197-223

Wei, K.C.J. and R. Chiang (1997). GMM and MLE Approaches for Estimation of Volatility and Regression Models when Daily Prices are Subjected to Price Limits. *Memo, Department of Finance, School of Business and Management,* The Hong Kong University of Science and Technology

Weiss, A. (1986). Asymptotic Theory for ARCH models: Estimation and Testing. *Econometric Theory*, 2, 107-131

Westerfield, R. (1977). The Distribution of Common Stock Price Changes: An Application of Transactions Time and Subordinated Stochastic Models. *Journal of Financial and Quantitative Analysis*, 12, 743-765

Whittle, P. (1953). Estimation and Information in Stationary Time Series. *Arkiv for Matematic,* 2, 423-434

Zakoian, J.M. (1990). Threshold Heteroskedastic Models. *CREST, INSEE*, Paris, Manuscript

Zarowin, P. (1990). Size, Seasonality and Stock Market Overreaction. *Journal of Financial and Quantitative Analysis*, 25, 113-125

Appendix

The ARCH models that have been presented

Table 2. The ARCH models that have been presented in Section 2.1. The reader who is interested in more information for an ARCH model should recur to the equation referred in the last column.

ARCH(q)	$\sigma_t^2 = a_0 + \sum_{i=1}^{q}\left(a_i \varepsilon_{t-i}^2\right)$	Engle (1982)	(2.3)
GARCH(p,q)	$\sigma_t^2 = a_0 + \sum_{i=1}^{q}\left(a_i \varepsilon_{t-i}^2\right) + \sum_{j=1}^{p}\left(b_j \sigma_{t-j}^2\right)$	Bollerslev (1986)	(2.6)
IGARCH(p,q)	$\sigma_t^2 = a_0 + \sum_{i=1}^{q}\left(a_i \varepsilon_{t-i}^2\right) + \sum_{j=1}^{p}\left(b_j \sigma_{t-j}^2\right)$, for $\sum_{i=1}^{q} a_i + \sum_{j=1}^{p} b_j = 1$	Engle and Bollerslev (1986)	(2.8)
EGARCH(p,q)	$\log\left(\sigma_t^2\right) = a_0 + \left(1 + \sum_{i=1}^{q} a_i L^i\right)\left(1 - \sum_{j=1}^{p} b_j L^j\right)^{-1} \left(\theta\left(\left\|\varepsilon_{t-1}/\sigma_{t-1}\right\| - E\left\|\varepsilon_{t-1}/\sigma_{t-1}\right\|\right) + \gamma\left(\varepsilon_{t-1}/\sigma_{t-1}\right)\right)$	Nelson (1991)	(2.13)
GJR(p,q)	$\sigma_t^2 = a_0 + \sum_{i=1}^{q}\left(a_i \varepsilon_{t-i}^2\right) + \sum_{i=1}^{q}\left(\gamma_i d\left(\varepsilon_{t-i} < 0\right)\varepsilon_{t-i}^2\right) + \sum_{j=1}^{p}\left(b_j \sigma_{t-j}^2\right)$	Glosten et al. (1993)	(2.14)
TGARCH(p,q)	$\sigma_t = a_0 + \sum_{i=1}^{q}\left(a_i \varepsilon_{t-i}^{+}\right) - \sum_{i=1}^{q}\left(\gamma_i \varepsilon_{t-i}^{-}\right) + \sum_{j=1}^{p}\left(b_j \sigma_{t-j}\right)$	Zakoian (1990)	(2.15)
AGARCH(p,q)	$\sigma_t = a_0 + \sum_{i=1}^{q} a_i \left\|\varepsilon_{t-i}\right\| + \sum_{j=1}^{p} b_j \sigma_{t-j}$	Taylor (1986) Schwert (1989a,b)	(2.16)
Log-GARCH(p,q)	$\log\left(\sigma_t^2\right) = a_0 + \sum_{i=1}^{q} a_i \log\left(\varepsilon_{t-i}^2\right) + \sum_{j=1}^{p} b_j \log\left(\sigma_{t-j}^2\right).$	Geweke (1986) Pantula (1986)	(2.17)
Stdev-ARCH(q)	$\sigma_t^2 = \left(a_0 + \sum_{i=1}^{q} a_i \left\|\varepsilon_{t-i}\right\|\right)^2$	Schwert (1990)	(2.18)
NARCH(p,q)	$\sigma_t^{\delta} = a_0 + \sum_{i=1}^{q} a_i \left\|\varepsilon_{t-i}^2\right\|^{\delta/2} + \sum_{j=1}^{p} b_j \sigma_{t-j}^{\delta}$	Higgins and Bera (1992)	(2.19)
AGARCH(p,q)	$\sigma_t^2 = a_0 + \sum_{i=1}^{q}\left(a_i \varepsilon_{t-i}^2 + \gamma_i \varepsilon_{t-i}\right) + \sum_{j=1}^{p} b_j \sigma_{t-j}^2$	Engle (1990)	(2.21)
NAGARCH(p,q)	$\sigma_t^2 = a_0 + \sum_{i=1}^{q} a_i \left(\varepsilon_{t-i} + \gamma_i \sigma_{t-i}\right)^2 + \sum_{j=1}^{p} b_j \sigma_{t-j}^2$	Engle and Ng (1993)	(2.22)
VGARCH(p,q)	$\sigma_t^2 = a_0 + \sum_{i=1}^{q} a_i \left(\varepsilon_{t-i}/\sigma_{t-i} + \gamma_i\right)^2 + \sum_{j=1}^{p} b_j \sigma_{t-j}^2$	Engle and Ng (1993)	(2.23)

APARCH(p,q)	$\sigma_t^{\delta} = a_0 + \sum_{i=1}^{q} a_i \left(\lvert \varepsilon_{t-1} \rvert - \gamma_i \varepsilon_{t-i} \right)^{\delta} + \sum_{j=1}^{p} b_j \sigma_{t-j}^{\delta}$	Ding et al. (1993)	(2.24)
GQARCH(p,q)	$\sigma_t^2 = a_0 + \sum_{i=1}^{q} a_i \varepsilon_{t-i}^2 + \sum_{i=1}^{q} \gamma_i \varepsilon_{t-i} + 2 \sum_{i=1}^{q} \sum_{j=i+1}^{q} a_{ij} \varepsilon_{t-i} \varepsilon_{t-j} + \sum_{j=1}^{p} b_j \sigma_{t-j}^2$	Sentana (1995)	(2.25)
GQTARCH(p,q)	$\sigma_t^2 = \omega + \sum_{i=1}^{q} \sum_{j=1}^{J} a_{ij} I_j \left(\varepsilon_{t-i} \right) + \sum_{i=1}^{p} b_j \sigma_{t-j}^2$	Gouriéroux and Monfort (1992)	(2.28)
VSARCH(p,q):	$\sigma_t^2 = \omega + \sum_{i=1}^{q} a_i \varepsilon_{t-i}^2 + \gamma S_{t-1} \frac{\varepsilon_{t-1}^2}{\sigma_{t-1}^2} + \sum_{j=1}^{p} b_j \sigma_{t-j}^2$	Fornari Mele (1995)	(2.29)
AVSARCH(p,q)	$\sigma_t^2 = a_0 + \sum_{i=1}^{q} a_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} b_j \sigma_{t-j}^2 + \gamma S_{t-1} \varepsilon_{t-1}^2 + \delta \left(\left(\varepsilon_{t-1}^2 / \sigma_{t-1}^2 \right) - k \right) S_{t-1}$	Fornari and Mele (1995)	(2.30)
LST-GARCH(p,q)	$\sigma_t^2 = a_0 + \sum_{i=1}^{q} \left(a_i + \gamma_i F\left(\varepsilon_{t-i} \right) \right) \varepsilon_{t-i}^2 + \sum_{j=1}^{p} b_j \sigma_{t-j}^2$ $F\left(\varepsilon_{t-i} \right) = \left(1 + \exp\left(-\theta \varepsilon_{t-i} \right) \right)^{-1} - 0.5$	Hagerud (1996)	(2.32)
EST-GARCH(p,q)	$\sigma_t^2 = a_0 + \sum_{i=1}^{q} \left(a_i + \gamma_i F\left(\varepsilon_{t-i} \right) \right) \varepsilon_{t-i}^2 + \sum_{j=1}^{p} b_j \sigma_{t-j}^2$ $F\left(\varepsilon_{t-i} \right) = 1 - \exp\left(-\theta \varepsilon_{t-i}^2 \right)$	Hagerud (1996)	(2.33)
GLST-GARCH(p,q)	$\sigma_t^2 = a_0 + \sum_{i=1}^{q} \left(a_i + \gamma_i F\left(\varepsilon_{t-i} \right) \right) \varepsilon_{t-i}^2 + \sum_{j=1}^{p} b_j \sigma_{t-j}^2$ $F\left(\varepsilon_{t-i} \right) = \frac{1 - \exp\left(-\theta \varepsilon_{t-i}^2 \right)}{1 + \exp\left(-\theta \left(\varepsilon_{t-i}^2 - c^2 \right) \right)}$	Lubrano (1998)	(2.34)
GEST-GARCH(p,q)	$\sigma_t^2 = a_0 + \sum_{i=1}^{q} \left(a_i + \gamma_i F\left(\varepsilon_{t-i} \right) \right) \varepsilon_{t-i}^2 + \sum_{j=1}^{p} b_j \sigma_{t-j}^2$ $F\left(\varepsilon_{t-i} \right) = 1 - \exp\left(-\theta \left(\varepsilon_{t-i} - c \right)^2 \right)$	Lubrano (1998)	(2.35)
CGARCH(1,1)	$\sigma_t^2 = q_t + a_1 \left(\varepsilon_{t-1}^2 - q_{t-1} \right) + b_1 \left(\sigma_{t-1}^2 - q_{t-1} \right)$ $q_t = a_0 + p q_{t-1} + \phi \left(\varepsilon_{t-1}^2 - \sigma_{t-1}^2 \right)$	Engle and Lee (1993)	(2.37)
ACGARCH(1,1)	$\sigma_t^2 = q_t + a_1 \left(\varepsilon_{t-1}^2 - q_{t-1} \right) + \gamma_1 \left(d\left(\varepsilon_{t-1} < 0 \right) \varepsilon_{t-1}^2 - 0.5 q_{t-1} \right) + b_1 \left(\sigma_{t-1}^2 - q_{t-1} \right)$ $q_t = a_0 + p q_{t-1} + \phi \left(\varepsilon_{t-1}^2 - \sigma_{t-1}^2 \right) + \gamma_2 \left(d\left(\varepsilon_{t-1} < 0 \right) \varepsilon_{t-1}^2 - 0.5 \sigma_{t-1}^2 \right)$	Engle and Lee (1993)	(2.39)
FIGARCH(p,d,q)	$\sigma_t^2 = a_0 + \left(1 - B(L) - \Phi(L) (1 - L)^d \right) \varepsilon_t^2 + B(L) \sigma_t^2$	Baillie et al. (1996)	(2.46)
FIEGARCH(p,d,q)	$\log\left(\sigma_t^2 \right) = a_0 + \Phi(L)^{-1} (1 - L)^{-d} \left(1 + A(L) \right) g\left(z_{t-1} \right)$	Bollerslev and Mikkelsen (1996)	(2.47)
FIAPARCH(p,d,q)	$\sigma_t^{\delta} = a_0 + \left(1 - \left(1 - B(L) \right)^{-1} \Phi(L) (1 - L)^{-d} \right) \left(\lvert \varepsilon_t \rvert - \gamma \varepsilon_t \right)^{\delta}$	Tse (1998)	(2.48)

ASYMM FIFGARCH(1,d,1)	$\sigma_t^{\lambda} = \frac{k}{1-\delta} + \left(1 - \frac{(1-\varphi L)(1-L)^d}{1-\delta L}\right) f^{v}(\varepsilon_t)\sigma_t^{\lambda}$ $f(\varepsilon_t) = \left\lvert \frac{\varepsilon_t}{\sigma_t} - b \right\rvert - c\left(\frac{\varepsilon_t}{\sigma_t} - b\right),$	Hwang (2001)	(2.49)
ASYMM FIFGARCH(1,d,1) modified	$(1-\varphi L)(1-L)^d \frac{\sigma_t^{\lambda} - 1}{\lambda} = \omega' + a(1+\psi L)\sigma_{t-1}^{\lambda}\left(f^{v}(z_{t-1}) - 1\right)$ $f\left(\frac{\varepsilon_t}{\sigma_t}\right) = \left\lvert \frac{\varepsilon_t}{\sigma_t} - b \right\rvert - c\left(\frac{\varepsilon_t}{\sigma_t} - b\right)$	Ruiz and Perez (2002)	(2.50)
R-GARCH(r,p,q)	$\sigma_t^2 = \sum_{i^*=1}^{r}\left(c_{i^*}\eta_{t-i^*}\right) + \sum_{i=1}^{q}\left(a_i\varepsilon_{t-i}^2\right) + \sum_{j=1}^{p}\left(b_j\sigma_{t-j}^2\right)$	Nowicka-Zagrajek and Weron (2001)	(2.51)
H-GARCH(p,n)	$\sigma_t^2 = a_0 + \sum_{i=1}^{n}\sum_{k=1}^{i} a_{ik}\left(\sum_{i^*=k}^{i}\varepsilon_{t-i^*}^2\right)^2 + \sum_{j=1}^{p}\left(b_j\sigma_{t-j}^2\right)$	Müller et al. (1997)	(2.52)