CHAPTER 1

1. ARCH models

1.1 Introduction

A crucial assumption in many statistical models is that of constant variance. Lately, a family of time series models has been developed relaxing the assumption of constant variance through time. This family is called Autoregressive Conditional Heteroskedasticity (ARCH) Models and was introduced by Engle (1982). These are conditional mean zero, serially uncorrelated stochastic processes with non-constant variances conditional on past, but invariant unconditional variances. The variance of the dependent variable is modeled as a function of past values of the dependent variable and independent, or exogenous variables. Autoregressive Conditional Heteroskedasticity (ARCH) models are specifically designed to model and forecast conditional variances. They been successfully applied in macroeconomic and financial time series in order to model and forecast volatility. Some of the areas where the ARCH models are widely used are: i) portfolio risk analysis, ii) option pricing, iii) time-varying confidence intervals forecasting. The aim is to obtain more accurate intervals of conditional mean by modeling the variance of the errors.

Our research proceeds as follows. Chapter 1 presents the most important regularities govern asset returns volatility and the incorporation of them in modeling both the conditional mean and conditional variance. The next Chapter provides the formulation of maximum likelihood estimators and their properties under the assumption of normality and under the absence of normality. In Chapter 3, we examine the dynamic structure of Greek Stock Market. Chapter 4 contains an empirical application of ARCH processes in Greek Stock Market. Chapter 5 deals with the conclusions of the research.

1.2 Autoregressive Conditional Heteroskedasticity Processes

Consider a stochastic process of interest $\{y_t(\theta_0)\}$ parametrized by the finite dimensional vector $\theta_0 \in \Theta \subseteq R^m$, where θ_0 denotes the true value, with conditional mean

$$\mu_t(\theta_0) = E(y_t \mid I_{t-1}) = E_{t-1}(y_t) \quad t=1,2,\ldots$$

I_{t-1} denotes any information available at time $t-1$ (Information Set at time $t-1$). Define the $\{u_t(\theta_0)\}$ process by

$$\{u_t(\theta_0)\} \equiv y_t - \{\mu_t(\theta_0)\} \quad t=1,2,\ldots$$

The $\{u_t(\theta_0)\}$ process is then defined to follow an ARCH model if the conditional mean equals zero,

$$E_{t-1}(u_t(\theta_0)) = 0,$$

but the conditional variance varies through time,

$$h_t(\theta_0) = Var_{t-1}(u_t(\theta_0)) = E_{t-1}(u_t^2(\theta_0)).$$

1.2.1 Modeling the conditional variance

Numerous parametric specifications for the time varying conditional variance have been proposed in the literature. The first model is the **ARCH(q)** model introduced by Engle (1982). The conditional variance is a linear function of the past q squared innovations[1]:

$$h_t = a_0 + \sum_{i=1}^{q} a_i u_{t-i}^2, \; a_0 > 0, \; a_i \geq 0, \; i=1,\ldots,q.$$

In empirical applications of ARCH(q) models a long lag length and a large number of parameters are often needed. Thus, Bollerslev (1986) generalized the ARCH(q) model and introduced the General Autoregressive Conditional Heteroskedasticity **GARCH(p,q)**

[1] The term "innovation" is used instead of the "residual" and expresses the unpredictable part of a financial series.

model. The conditional variance is a linear function of the past q squared innovations and the past p conditional variances:

$$h_t = a_0 + \sum_{i=1}^{q} a_i u_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j}\ ,\ a_0 > 0,\ a_i \geq 0,\ \text{i=1,…,q},$$

$$\beta_j \geq 0,\ \text{j=1,…,p.}$$

Note that the model is covariance stationary if and only if $\sum_{i=1}^{q} a_i + \sum_{j=1}^{p} \beta_j < 1$. In this case, the unconditional variance of u_t is given by:

$$E(u_t^2) = h = a_0 \left[1 - \sum_{i=1}^{q} a_i - \sum_{j=1}^{p} \beta_j \right]^{-1}.$$

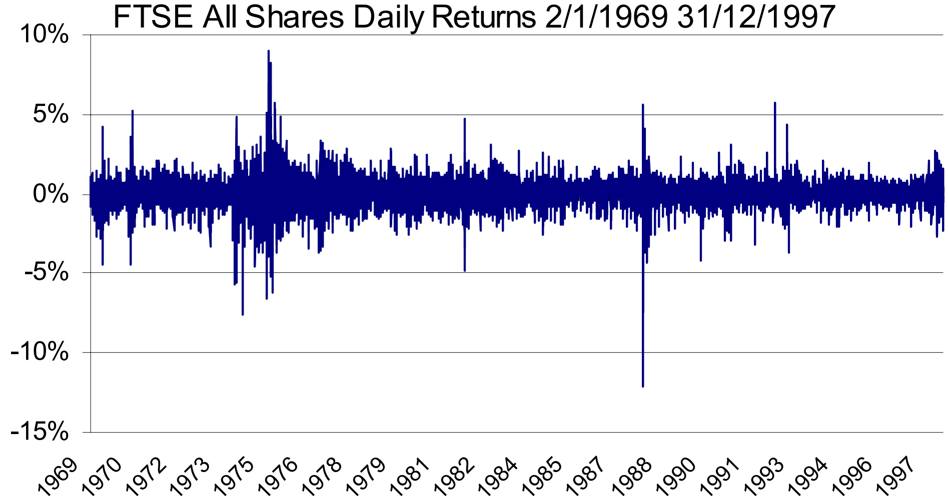
In empirical investigations the estimate of $\sum_{i=1}^{q} a_i + \sum_{j=1}^{p} \beta_j$ is very close to unity. Engle and Bollerslev (1986) referred to a model satisfying $\sum_{i=1}^{q} a_i + \sum_{j=1}^{p} \beta_j = 1$ as an integrated GARCH process, denoted **IGARCH(p,q)**. Under an IGARCH process the unconditional variance of u_t is infinite, so neither u_t not u_t^2 satisfies the definition of a covariance stationary process.

GARCH models are suitable to capture some characteristics of financial markets. They elegantly capture the volatility clustering in asset returns 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…". The volatility clustering phenomenon is apparent when asset returns are plotted through time. Figure 1.1 plots the daily returns[1] on the English FTSE All Shares stock index from 1969 to 1997 (we are grateful to GrStocks.com for providing the data).

[1] The returns are expressed in percent and are continuously compounded.

Figure 1.1

FTSE All Shares Daily Returns 2/1/1969 31/12/1997



It is clear from visual inspection of the figure that the returns are not independent identically distributed through time. Volatility was clearly higher during the 1970's than during the 1990's. Large changes tend to be followed by large changes around 1975, of either sign, and small changes tend to be followed by small changes during the last years. The structure of a GARCH model imposes an important limitation. GARCH models assume that only the magnitude and not the positivity or negativity of innovations determines the feature of h_t because h_t is a function of lagged h_t and lagged u_t^2 and so is invariant to changes in the algebraic sign of the u_t^2's.

On the other hand, asset returns tend to be leptokurtic (heavily tailed). For example, the kurtosis for the daily returns on the FTSE All Shares is 13,02. Denote as

$$z_t(\theta_0) \equiv u_t(\theta_0) \big/ \sqrt{h_t(\theta_0)}$$

the standardized process, it will have conditional mean zero and time invariant conditional variance unity. If the conditional distribution for z_t is furthermore assumed to be time invariant with finite fourth moment, then the unconditional distribution for u_t will have fatter tails than the distribution for z_t. For instance, for the ARCH(1) model with conditionally normally distributed errors, $E(u_t^4)/E^2(u_t^2) = 3(1-a_1^2)/(1-3a_1^2)$, if $3a_1^2 < 1$, and $E(u_t^4)/E^2(u_t^2) = \infty$ otherwise; both of which exceed the normal value of three.

Financial markets are characterized by the so called "leverage effect", first noted by Black (1976). The "leverage effect" refers to the tendency for the changes in the stock prices to be negatively correlated with changes in stock volatility. I.e. volatility tends to rise in response to "bad news" (returns lower than expected) and fall in response to "good news" (returns higher than expected).

Exponential GARCH MODEL

Nelson (1991), proposed the following model for the evolution of the conditional variance of u_t:

$$\log(h_t) = a_0 + \sum_{j=1}^{\infty} \pi_j \left\{ \left| \frac{u_{t-j}}{\sqrt{h_{t-j}}} \right| - E\left| \frac{u_{t-j}}{\sqrt{h_{t-j}}} \right| + \delta \frac{u_{t-j}}{\sqrt{h_{t-j}}} \right\}$$

This model is referred to as exponential GARCH, or EGARCH. In this model, h_t depends on both the magnitude and the sign of lagged residuals. The δ parameter allows for the asymmetric effect. If $\delta = 0$ then a positive surprise has the same effect on volatility as a negative surprise. If $-1 < \delta < 0$, a positive surprise increases volatility less than a negative surprise. If $\delta < -1$, a positive surprise actually reduces volatility while a negative surprise increases volatility. For $\delta < 0$ the "leverage effect" exists. Since EGARCH describes the log of h_t, the h_t will be positive regardless of whether the π_j coefficients are positive. Thus, in contrast to the GARCH model, no restrictions need to be imposed on the model for estimation. Theorem 2.1 in Nelson (1991) implies that $\log h_t, h_t$ and u_t are all strictly stationary, provided that $\sum_{j=1}^{\infty} \pi_j^2 < \infty$. We can express the infinite moving average representation of the model as the ratio of two finite order polynomials. Thus, an ARMA process provides a simpler parameterization of the form.

We denote it as **EGARCH(p,q)**:

$$\log(h_t) = a_0 + \sum_{j=1}^{p}\left\{ a_j \left|\frac{u_{t-j}}{\sqrt{h_{t-j}}}\right| - a_j E\left|\frac{u_{t-j}}{\sqrt{h_{t-j}}}\right| + \delta_j \frac{u_{t-j}}{\sqrt{h_{t-j}}}\right\} + \sum_{i=1}^{q}\left(\beta_i \log(h_{t-i})\right).$$

The EGARCH model can be estimated by the maximum likelihood method by specifying a density for z_t. Nelson proposed as density function for the standardized process $z_t \equiv u_t / \sqrt{h_t}$ the generalized error distribution[1] (Harvey (1981), Box and Tiao (1973)) normalized to have zero mean and unit variance:

$$f(z_t) = \frac{v \cdot e^{-2^{-1}\left|\frac{z_t}{\lambda}\right|^{v}}}{\lambda \cdot 2^{\frac{v+1}{v}}\Gamma\left(\frac{1}{v}\right)}$$

where $\Gamma(\cdot)$ denotes the gamma function, λ is a constant given by $\lambda \equiv \left(\frac{2^{-2v^{-1}}\Gamma\left(\frac{1}{v}\right)}{\Gamma\left(\frac{3}{v}\right)}\right)^{\frac{1}{2}}$ and v is a positive parameter governing the thickness of the tails. Note that for $v = 2$, the constant λ is equal to 1 and the generalized error distribution reduces to the standard normal density. If $v < 2$ the density has thicker tails than the normal whereas for $v > 2$ it has thinner tails. I.e. for $v = 1$, z_t has a double exponential distribution whereas for $v = \infty$, z_t is uniformly distributed on the interval $\left[-\sqrt{3}, \sqrt{3}\right]$.

The family of ARCH models is remarkably rich. Another route for introducing asymmetric effects is to set:

$$\sqrt{h_t} = a_0 + \sum_{i=1}^{q}\left[a_i^{+} I(u_{t-i} > 0)\left|u_{t-1}\right| + a_i^{-} I(u_{t-i} \le 0)\left|u_{t-1}\right|\right] + \sum_{j=1}^{p}\beta_j \sqrt{h_{t-j}},$$

where $I(\cdot)$ denotes the indicator function[2]. The model introduced by Zakoian (1990) is called Threshold ARCH or **TARCH(p,q)**.

Glosten, Jagannathan and Runkle (1993) introduced the **GJR(p,q)** model with the following form:

---

[1] Box and Tiao call the GED the exponential power distribution.

[2] $I(u_{t-i} > 0) = 1$ if $u_{t-i} > 0$, otherwise zero. $I(u_{t-i} \le 0) = 1$ if $u_{t-i} \le 0$, otherwise zero.

$$h_t = a_0 + \sum_{i=1}^{q} a_i u_{t-i}^2 + \delta_1 I(u_{t-1} < 0) u_{t-1}^2 + \sum_{j=1}^{p} \beta_j h_{t-j} ,$$

where $I(\cdot)$ denotes the indicator function. The "leverage effect" is supported if $\delta_1 > 0$. "Good news" has got an impact of a_i and "bad news" has got an impact of $a_i + \delta_1$.

Engle (1990), proposed the Asymmetric ARCH or **AARCH(p,q)** model:

$$h_t = a_0 + \sum_{i=1}^{q} \left(a_i u_{t-i}^2 + \delta_i u_{t-i} \right) + \sum_{j=1}^{p} \beta_j h_{t-j} ,$$

where a negative value of δ_i means that positive returns increase volatility less than negative returns.

Taylor (1986) modeled the conditional standard deviation function instead of conditional variance. Schwert (1989) modeled the conditional standard deviation as a linear function of lagged absolute residuals.

The **Taylor/Schwert GARCH(p,q)** model is defined as

$$h_t^{1/2} = a_0 + \sum_{i=1}^{q} a_i |u_{t-1}| + \sum_{j=1}^{p} \beta_j h_{t-j}^{1/2} .$$

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

$$h_t^{\gamma/2} = a_0 + \sum_{i=1}^{q} a_i |u_{t-1}|^{\gamma/2} + \sum_{j=1}^{p} \beta_j h_{t-j}^{\gamma/2} .$$

Geweke (1986) and Pantula (1986) introduced the **log-ARCH(p,q)** model:

$$\log(h_t) = a_0 + \sum_{i=1}^{q} a_i \log\left(u_{t-i}^2\right) + \sum_{j=1}^{p} \beta_j \log\left(h_{t-j}\right).$$

Sentana (1995) introduced the Quadratic ARCH or **QARCH(p,q)** model of the form

$$h_t = a_0 + \sum_{i=1}^{q} a_i u_{t-i}^2 + \sum_{i=1}^{q} \delta_i u_{t-i} + 2 \sum_{i=1}^{q} \sum_{j=i+1}^{q} \delta_{ij} u_{t-i} u_{t-j} + \sum_{j=1}^{p} \beta_j h_{j-1} .$$

Ding, Granger and Engle (1993) introduced the Asymmetric Power ARCH or **APARCH(p,q)** model:

$$h_t^{\gamma/2} = a_0 + \sum_{i=1}^{q} a_i \left(|u_{t-1}| - \delta_i u_{t-i} \right)^{\gamma/2} + \sum_{j=1}^{p} \beta_j h_{t-j}^{\gamma/2} ,$$

which includes seven ARCH models as special cases[1]. Ding, Granger and Engle (1993) estimate the Standard & Poor's 500 (hereafter S&P 500) returns by the APARCH(1,1) model and the estimated power $\gamma / 2$ for the conditional heteroskedasticity function is 1.43, which is significantly different from 1 (Taylor/Schwert model) or 2 (GARCH model).

Non-trading periods

Information that accumulates when financial markets are closed is reflected in prices after the markets reopen. If, for example, information accumulates at a constant rate over calendar time, then the variance of returns over the period from the Friday close to the Monday close should be three times the variance from the Monday close to the Tuesday close. Fama (1965) and French and Roll (1986) have found, however, that information accumulates more slowly when the markets are closed than when they are open. Variances are higher following weekends and holidays than on other days, but not nearly by as much as would be expected if the news arrival rate were constant. For instance, using data on daily returns across all NYSE stocks from 1963 to 1982, French and Roll (1986) found that volatility is 70 times higher per hour on average when the market is open than when it is closed. Baillie and Bollerslev (1989) report qualitatively similar results for foreign exchange rates.

1.2.2 Modeling the conditional mean

The conditional mean $\mu_t(\theta_0) = E_{t-1}(y_t)$ should be modeled in order to incorporate information from empirical regularities of asset returns.

Non-synchronous trading

According to efficient market theory, the stock market returns themselves contain little serial correlation. Moreover, when high frequency data is used, the non-synchronous trading in the stocks making up an index induces positive first order serial correlation in

[1] ARCH, GARCH, Taylor/Schwert GARCH, GJR, TARCH, NARCH and logARCH

the return series. To control this Scholes and Williams (1977) suggested a first order moving average form, while Lo and Mackinlay (1988) suggested a first order autoregressive 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”.

Risk return tradeoff

Many theories in finance are dealt with the tradeoff between the expected returns and variance, or the covariance among the returns. According to the Capital Asset Pricing Model (CAPM) the excess returns[1] on all risky assets are proportional to the non-diversifiable risk as measured by the covariances with the market portfolio.

The CAMP in its traditional form is as follows:

$$E(R_i) = R_f + \left[E(R_m) - R_f\right]\frac{\sigma_{im}}{\sigma_m^2}.$$

The expected rate of return on i asset $E(R_i)$ is equal to the risk free rate of return R_f plus a risk premium. The risk premium is the price of risk multiplied by the quantity of risk. The price of risk is the difference between the expected rate of return on the market portfolio and the risk free rate of return. The quantity of risk is often called beta and is:

$$\beta_i = \frac{\sigma_{im}}{\sigma_m^2} = \frac{Cov(R_i, R_m)}{Var(R_m)}.$$

The CAPM is based on the Capital Market Line (CML). The CML is the linear efficient set[2], which is the same for all the investors under the assumption they have homogeneous beliefs. The equation for the CML is:

$$E(R_p) = R_f + \frac{\left\lfloor E(R_m) - R_f \right\rfloor}{\sigma(R_m)}\sigma(R_p).$$

[1] Asset return minus the risk free interest rate. As an approximation to the risk free interest rate we usually use the three month Treasury Bill return.

[2] The efficient set is the set of mean-variance choices from the investment opportunity set where for a given variance no other investment opportunity offers a higher mean return.

The expected return for an efficient portfolio p of assets, $E(R_p)$, is equal to the risk free rate of return, R_f, plus a slope, $\frac{\lfloor E(R_m) - R_f \rfloor}{\sigma(R_m)}$, times the variance of returns on the efficient portfolio p. Thus, the CML provides a simple linear relationship between the risk and return for efficient portfolios of assets.[1] Merton (1973) in Intertemporal Capital Asset Pricing Theory showed that the expected excess return on the market portfolio is linear in its conditional variance.

The ARCH in mean or ARCH-M model, introduced by Engle et al. (1987), was designed to capture such relationships. In the ARCH-M model the conditional mean is an explicit function of the conditional variance:

$$\mu_t(\theta) = g[h_t(\theta), \theta],$$

where the derivative of the $g(.,.)$ function with respect to the first element is non-zero. The most commonly employed specifications of the ARCH-M model postulate a linear relationship in h_t or $h_t^{1/2}$, e.g. $g[h_t(\theta), \theta] = \mu_0 + \mu_1 h_t$. A positive as well as a negative relationship between risk and return could be consistent with the financial theory. We expect a positive relationship if we assume a rational risk averse investor who requires a larger risk premium during times the payoff of the security is riskier. But we expect a negative relationship under the assumption that during relatively riskier periods the investors may want to save more. In applied researches, there is evidence for both relationships. French, Schwert and Stambaugh (1987) found positive risk return tradeoff for the excess returns on the Standard & Poor's composite portfolio although statistically significant not in all the periods. Bollerslev, Engle and Nelson (1994) found positive, not always statistically significant, relationship for the returns on Dow Jones, Standard 90 and S&P 500. Nelson (1991) found negative but insignificant relationship for the excess returns on the CRSP (Center for Research in Security Prices) value weighted market index.

Volatility and Serial Correlation

LeBaron (1992) found a strong inverse relation between volatility and serial correlation

[1] For more information about CAPM, CML and efficient portfolios see Copeland and Weston (1992).

for Standard & Poor's, CRSP value weighted index, Dow Jones and IBM returns. He introduced the Exponential Autoregressive GARCH model or EXP-GARCH in which the conditional mean is a non-linear function of conditional variance:

$$\mu_t(\theta) = \mu_2 e^{\frac{-h_t}{\mu_3}} y_{t-1} ,$$

As LeBaron stated, it is difficult to estimate μ_3 in conjunction with μ_2 when using a gradient type of algorithm. For this reason, μ_2 is set to the sample variance of the series. LeBaron found that μ_2 is significantly negative and remarkably robust to the choice of sample period, market index, measurement interval and volatility measure.