Chapter 2

Models for Financial Time Series

2.1 Introduction

Financial time series are well known for their uncertainty, especially the irregularity in the behaviour of certain financial indices such as stock prices, exchange or interest rates, government bond prices, yield of treasury bills and so on, that are prone to time dependent variability. Such variability, otherwise known as volatility can generate very high frequency series of variables which are stochastic in nature, the dynamics of which can best be described by means of stochastic models. As a result of the added uncertainty, statistical theory and methods play an important role in financial time series analysis.

There are two main objectives of investigating financial time series. First, it is important to understand how prices behave. The variance of the time series is particularly relevant. Tomorrow’s price is uncertain and it must therefore be described
by a probability distribution. This means that statistical methods are the natural way to investigate prices. Usually one builds a model, which is a detailed description of how successive prices are determined. The second objective is to use our knowledge of price behaviour to reduce risk or take better decisions. Time series models may for instance be used for forecasting, option pricing and risk management. This motivates more and more statisticians and econometricians to devote themselves to the development of new (or refined) time series models and methods.

Many finance problems involve the arrival of events such as prices or trades in irregular time intervals, a new direction of modelling is necessary to explain the properties of such data. The durations between market activities such as trades, quotes, etc. provide useful information on the underlying assets while analysing financial time series. Hence it is important to model the dynamic behaviour of such durations in finance.

The objective of this chapter is to understand various aspects of financial time series and list some of the important financial time series models and their useful characteristics. In the next section, we address some of the stylized facts of financial time series which play important role in volatility modelling. Section 2.3 introduces models for volatility and basic properties. In Section 2.4 we discuss about the conditional duration models in finance.
2.2 Stylized facts of Financial Time Series

Financial time series analysis is concerned with the theory and practice of asset valuation over time. One of the objectives of analysing financial time series is to model the volatility and forecast its future values. The volatility is measured in terms of the conditional variance of the random variables involved. The conditional variances in the case of financial time series are not constants. They may be functions of some known or unknown factors. This leads to the introduction of conditional heteroscedastic models for analysing financial time series. In financial markets, the data on price \( P_t \) of an asset at time \( t \) is available at different time points. However, in financial studies, the experts suggest that the series of returns be used for analysis instead of the actual price series, see Tsay (2005). For a given series of prices \( \{P_t\} \), the corresponding series of returns is defined by

\[
R_t = \frac{P_t - P_{t-1}}{P_{t-1}} = \frac{P_t}{P_{t-1}} - 1, \quad t = 1, 2, \ldots
\]

The advantages of using the return series are, 1) for an investor, the return series is a scale free summary of the investment opportunity, 2) the return series are easier to handle than the price series because of their attractive statistical properties. Further consideration of the attractive statistical properties, suggested that, the log-return series defined by \( r_t = \log(P_t/P_{t-1}) \) is more suitable for analysing the stochastic nature of the market behaviour. Hence, we focus our attention on the modelling and analysis of the log-return series in this thesis and we refer \( \{r_t = \log(P_t/P_{t-1}), t = 1, 2, \ldots\} \) as financial time series.
Empirical studies on financial time series (See Mandebrot (1963) and Fama (1965)) show that the series \( \{r_t\} \) defined above is characterized by the properties such as

1. Absence of autocorrelation in \( \{r_t\} \).
2. Significant serial correlation in \( \{r_t^2\} \).
3. The marginal distribution \( \{r_t\} \) is heavy-tailed.
4. Conditional variance of \( r_t \) given the past is not constant.
5. Volatility tends to form clusters, i.e., after a large (small) price change (positive or negative) a large (small) price change tends to occur. This attribute is called volatility clustering.

To get an intuitive feel of these stylized facts, a typical example is shown in Figure 2.1, where Bombay Stock Exchange (BSE) opening index during July 02, 2007 to May 13, 2016 is plotted.

**Figure 2.1:** Time series plot of BSE index and returns
Right panel in Figure 2.1 plots the time series of returns of the indices under study. Time series plot of indices are clearly non-stationary, however daily returns are stationary.

Summary statistics for daily index returns $r_t$ are provided in Table 2.1. These statistics are used in the discussion of some stylized facts related to the probability density function of the return series.

<table>
<thead>
<tr>
<th>Statistics</th>
<th>BSE index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observations</td>
<td>2185</td>
</tr>
<tr>
<td>Mean</td>
<td>0.0003</td>
</tr>
<tr>
<td>Median</td>
<td>0.0004</td>
</tr>
<tr>
<td>Maximum</td>
<td>0.1205</td>
</tr>
<tr>
<td>Minimum</td>
<td>-0.1138</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>0.0166</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.3853</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>10.2901</td>
</tr>
</tbody>
</table>

Table 2.1: Summary statistics for BSE log-returns

As seen in Table 2.1, BSE index returns have excess kurtosis well above 3 indicates leptokurtic and fat tails of returns. The ACF of returns and squared returns are plotted in Figure 2.2. While the autocorrelation of returns are all close to zero, autocorrelation of squared returns are positive and significantly larger than zero. Since the autocorrelation is positive, it can be concluded, that small (positive or negative) returns are followed by small returns and large returns follow large ones again.

Figure 2.3 compares histogram of BSE index return with approximate normal density. It is clear from the figure that the empirical distribution of daily returns does
Chapter 2. Models for Financial time series

Figure 2.2: ACF of returns and squared returns of BSE index

not resemble a Gaussian distribution. The peak around zero appears clearly, but the thickness of the tails is more difficult to visualize.

Figure 2.3: Histogram of BSE index return and normal approximation
2.3 Models for Volatility

The models described in the previous chapter are often very useful in modelling time series in general. However, they have the assumption of constant error variance. As a result the conditional variance of the observation at any time given the past will remain a constant, a situation referred to as homoscedasticity. This is considered to be unrealistic in many areas of economics and finance as the conditional variances are non-constants. Therefore, two prominent classes of models have been developed by researchers which capture the time-varying autocorrelated volatility process: the autoregressive conditional heteroscedastic (ARCH) model, introduced by Engle (1982), assumes that the conditional variances are some functions of the squares of the past returns and are referred to as the observation driven models. Another class of models to study the price changes is the SV models introduced by Taylor (1986), where the conditional variance at time $t$ is assumed to be a stochastic process in terms of some latent variables, which are referred to as the parameter driven models.

2.3.1 Autoregressive Conditional Heteroscedastic Models

The ARCH model introduced by Engle (1982) was a first attempt in econometrics to capture volatility clustering in time series data. In particular, Engle (1982) used conditional variance to characterize volatility and postulated a dynamic model for conditional variance. We will discuss the properties and some generalizations of the ARCH model in subsequent sections; for a comprehensive review of this class of
models we refer to Bollerslev et al. (1992). 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. Specifically, an ARCH($p$) model for $\{r_t\}$ is defined by

$$r_t = \sqrt{h_t} \varepsilon_t, \quad h_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i r_{t-i}^2,$$

where $\{\varepsilon_t\}$ is a sequence of independent and identically distributed random variables with mean zero and variance 1, $\alpha_0 > 0$, and $\alpha_i \geq 0$ for $i > 0$. If $\{\varepsilon_t\}$ has standardized Gaussian distribution conditional on $h_t$, $r_t$ follows normal with mean 0 and variance $h_t$. The Gaussian assumption of $\varepsilon_t$ is not critical. We can relax it and allow for more heavy-tailed distributions, such as the Student’s $t$-distribution, as is typically required in finance. Now we describe the properties of a first order ARCH model in detail.

**ARCH(1) model and properties:**

The structure of the ARCH model implies that the conditional variance $h_t$ of $r_t$, evolves according to the most recent realizations of $r_t^2$ analogous to an AR(1) model. Large past squared shocks imply a large conditional variance for $r_t$. As a consequence, $r_t$ tends to assume a large value which in turn implies that a large shock tends to be followed by another large shock. To understand the ARCH models, let us now take a closer look at the ARCH(1) model

$$r_t = \sqrt{h_t} \varepsilon_t, \quad h_t = \alpha_0 + \alpha_1 r_{t-1}^2,$$
where $\alpha_0 > 0$ and $\alpha_1 \geq 0$.

1. The unconditional mean of $r_t$ is zero, since

$$E(r_t) = E(E(r_t|r_{t-1})) = E\left(\sqrt{h_t}E(\varepsilon_t)\right) = 0.$$ 

2. The conditional variance of $r_t$ is

$$E\left(r^2_t|r_{t-1}\right) = E(h_t \varepsilon^2_t|r_{t-1}) = h_t E(\varepsilon^2_t|r_{t-1}) = h_t = \alpha_0 + \alpha_1 r^2_{t-1}.$$ 

3. The unconditional variance of $r_t$ is

$$Var\left(r_t\right) = E\left(r^2_t\right) = E\left(E\left(r^2_t|r_{t-1}\right)\right)$$

$$= E\left(\alpha_0 + \alpha_1 r^2_{t-1}\right)$$

$$= \alpha_0 + \alpha_1 E\left(r^2_{t-1}\right)$$

$$= \frac{\alpha_0}{1 - \alpha_1}.$$ 

4. Assuming that the fourth moment of $r_t$ are finite, the Kurtosis $K$ of $r_t$, is given by

$$K = \frac{E(r^4_t)}{E(r^2_t)^2} = \frac{1 - \alpha_1^2}{1 - 3\alpha_1^2} > 3,$$

provided $\alpha_1^2 < 1/3$.

The ARCH model with a conditionally normally distributed $r_t$ leads to heavy tails in the unconditional distribution. In other words, the excess kurtosis of
$r_t$ is positive and the tail of the distribution of $r_t$ is heavier than that of the normal distribution.

5. The autocovariance of $r_t$ is defined by

$$Cov(r_t, r_{t-k}) = E(r_t r_{t-k}) - E(r_t) E(r_{t-k})$$

$$= E(r_t r_{t-k}) = E\left(\sqrt{h_t} \sqrt{h_{t-k}}\right) E(\varepsilon_t \varepsilon_{t-k}) = 0.$$ 

Then the autocorrelation function of $r_t$ is zero. The ACF of $\{r^2_t\}$ is $\rho_{r^2_t}(k) = \alpha_1^k$ and notice that $\rho_{r^2_t}(k) \geq 0$ for all $k$, a result which is common to all linear ARCH models.

Thus, the ARCH(1) process has a mean of zero, a constant unconditional variance, and a time varying conditional variance. The $\{r_t\}$ is stationary process for which $0 \leq \alpha_1 < 1$ is satisfied, since the variance of $r_t$ must be positive. These properties continue to hold for general ARCH models, but the formulas become more complicated for higher order ARCH models.

### 2.3.2 Generalized ARCH (GARCH) Models

The GARCH model is an extension of Engle’s work by Bollerslev (1986) that allows the conditional variance to depend on the previous conditional variances and the squares of previous returns. The possibility that estimated parameters in ARCH model do not satisfy the stationarity condition increases with lag. Thus GARCH
model is an alternative to ARCH model. The GARCH \((p,q)\) is defined by

\[
r_t = \sqrt{h_t} \varepsilon_t, \quad h_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i r_{t-i}^2 + \sum_{j=1}^{q} \beta_j h_{t-j},
\]  

(2.3)

where \{\varepsilon_t\} is a sequence of independent and identically distributed random variables with mean 0 and variance 1; \{\varepsilon_t\} is assumed to be independent of \{h_{t-i}, i \geq 1\}. \(\alpha_0, \alpha_i, \text{ and } \beta_j\) are unknown parameters satisfying \(\alpha_0 > 0, \alpha_i \geq 0, \beta_j \geq 0,\) and \(\sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i) < 1.\) The constraint on \(\alpha_i + \beta_i\) implies that the unconditional variance of \(r_t\) is finite, whereas its conditional variance \(h_t\) evolves over time. As before, \(\varepsilon_t\) is assumed to be a standard normal distribution.

**GARCH (1,1) model and properties:**

Let us now consider the GARCH (1,1) model, which is the most popular one for modelling asset-return volatility. We represent this model as

\[
r_t = \sqrt{h_t} \varepsilon_t, \quad h_t = \alpha_0 + \alpha_1 r_{t-1}^2 + \beta_1 h_{t-1},
\]  

(2.4)

where \(\varepsilon_t \sim N(0,1)\) and \(0 \leq \alpha_1, \beta_1 < 1, \alpha_1 + \beta_1 < 1.\)

1. The unconditional mean of \(r_t\) is zero, since

\[
E(r_t) = E(E(r_t|r_{t-1})) = E\left(\sqrt{h_t}E(\varepsilon_t)\right) = 0.
\]
2. The conditional variance of $r_t$ is

$$E (r_t^2 | r_{t-1}) = E (h_t \varepsilon_t^2 | r_{t-1})$$
$$= h_t E (\varepsilon_t^2 | r_{t-1})$$
$$= h_t = \alpha_0 + \alpha_1 r_{t-1}^2 + \beta_1 h_{t-1}.$$ 

3. The unconditional variance of $r_t$ is

$$Var (r_t) = E (r_t^2) = E (E (r_t^2 | r_{t-1})) = E \left( \alpha_0 + \alpha_1 r_{t-1}^2 + \beta_1 h_{t-1} \right)$$
$$= \alpha_0 + \alpha_1 E (r_{t-1}^2) + \beta_1 E (h_{t-1})$$

Under stationarity we get

$$Var(r_t) = \frac{\alpha_0}{1 - (\alpha_1 + \beta_1)}.$$ 

4. The Kurtosis of $r_t$, $K$, is given by

$$K = 3 \left[1 - (\alpha_1 + \beta_1)^2\right] \frac{1}{1 - (\alpha_1 + \beta_1)^2 - 2\alpha_1^2} > 3.$$ 

Consequently, similar to ARCH models, the tail of the marginal distribution of GARCH(1,1) process is heavier than that of a normal distribution if $1 - 2\alpha_1^2 - (\alpha_1 + \beta_1)^2 > 0$.

5. The ACF of $\{r_t\}$ is zero and the ACF of $\{r_t^2\}$ is given by

$$\rho_{r_t^2}(k) = (\alpha_1 + \beta_1)^{k-1} \frac{\alpha_1 (1 - \alpha_1 \beta_1 - \beta_1^2)}{1 - 2\alpha_1 \beta_1 - \beta_1^2}, \quad k = 1, 2, ....$$
2.3.3 Stochastic Volatility Models

This is a class of parameter driven model in which the volatility at time $t$ is described as a latent stochastic process, such as the one generated by an autoregressive model. An appealing feature of the SV model is its close relationship to financial economic theories. The univariate SV model proposed by Taylor (1986) is given by,

$$r_t = \varepsilon_t \exp(h_t/2), \quad h_t = \alpha + \rho h_{t-1} + \eta_t; \quad |\rho| < 1, \ t = 1, 2, ..., \quad (2.5)$$

where $\varepsilon_t$ and $\eta_t$ are two independent Gaussian white noises, with variances 1 and $\sigma^2_\eta$, respectively. Due to the Gaussianity of $\eta_t$, this model is called a log-normal SV model. Its major properties are discussed in Taylor (1986, 1994).

As $\eta_t$ is Gaussian, $\{h_t\}$ is a Gaussian autoregressive process. It will be (strictly and covariance) stationary if $|\rho| < 1$ with:

$$\mu_h = E(h_t) = \frac{\alpha}{1 - \rho},$$

$$\sigma^2_h = V(h_t) = \frac{\sigma^2_\eta}{1 - \rho^2}.$$

As $\{\varepsilon_t\}$ is always stationary, $\{r_t\}$ is stationary if and only if $\{h_t\}$ is stationary, $r_t$ being the product of two stationary process. All odd moments of $r_t$ vanish and using the property of log-normal distribution, all the even moments of $r_t$ can be obtained if $h_t$ is stationary. In particular the kurtosis is

$$K = \frac{E(r_t^4)}{E(r_t^2)^2} = 3 \exp\left(\sigma^2_h\right) \geq 3,$$
which shows that the SV model has fatter tails than the corresponding normal distribution. The dynamic properties of $r_t$ are easy to find. First, as $\{\varepsilon_t\}$ is independent and identically distributed, $\{r_t\}$ is a martingale difference and is a white noise if $|\rho| < 1$. As $h_t$ is a Gaussian AR(1),

$$
Cov \left( r_t^2, r_{t-k}^2 \right) = E \left( r_t^2 r_{t-k}^2 \right) - E \left( r_t^2 \right) E \left( r_{t-k}^2 \right)
$$

$$
= E \left( \exp (h_t + h_{t-k}) \right) - (E \left( \exp (h_t) \right))^2
$$

$$
= \exp \left( 2\mu_h + \sigma_h^2 \right) \left( \exp \left( \sigma_h^2 \rho^k \right) - 1 \right),
$$

and so

$$
\rho_{r_t^2}(k) = \frac{Cov \left( r_t^2, r_{t-k}^2 \right)}{V \left( r_t^2 \right)} = \frac{\exp \left( \sigma_h^2 \rho^k \right) - 1}{3 \exp (\sigma_h^2) - 1} \frac{\exp \left( \sigma_h^2 \right) - 1}{3 \exp (\sigma_h^2) - 1} \rho^k.
$$

Note that if $\rho < 0$, $\rho_{r_t^2}(k)$ can be negative, unlike ARCH models. This resembles the autocorrelation function of an ARMA(1,1) process. Thus the SV model behaves in a manner similar to the GARCH(1,1) model. Finally, note that there is no need for non-negativity constraints or for bounded kurtosis constraints on the coefficients. This is a great advantage with respect to GARCH models. A review of the properties of SV models may be found in Taylor (1994) and Tsay (2005).

Despite theoretical advantages, the SV models have not been popular as the ARCH models in practical applications. The main reason is that the likelihood function for the SV model is not easy to evaluate unlike in the case of ARCH models. A variety of estimation procedures have been proposed to overcome this difficulty, including, for example, the Generalized Method of Moments (GMM) used by Melino and Turnbull
(1990), the Quasi Maximum Likelihood (QML) approach followed by Harvey et al. (1994) and Ruiz (1994), the Efficient Method of Moments (EMM) applied by Gallant et al. (1997), and Markov Chain Monte Carlo (MCMC) procedures used by Jacquier et al. (1994) and Kim et al. (1998). For a survey of these estimation procedures, one can refer Ghysels et al. (1996), Broto and Ruiz (2004) and Bauwens et al. (2012).

2.4 Models for Durations

The statistical analysis of sequence of durations between events is well studied in the area of point processes which includes the renewal process as a special case. In the analysis of financial time series the durations between market activities such as trades, quotes, etc. provide important information on the underlying asset. These durations are irregularly time-spaced and they form a sequence of random variables. As a result, the number of financial transactions taken place in any interval forms a point process called financial point process. For a specified asset, longer durations indicate lack of trading activities, which in turn signify a period of no new information. On the other hand arrival of new information often results in heavy trading and hence leads to shorter durations. The dynamic behaviour of durations thus contains useful information about market activities. Furthermore, since financial markets typically take a period of time to uncover the effect of new information, active trading is likely to persist for a period of time, resulting in clusters of short durations. Consequently, durations might exhibit characteristics similar to those of asset volatility, which is an important aspect of financial time series. Such features may be captured in alternative ways through different dynamic
models based on either duration or intensity representation of a point process. These considerations motivated Engle and Russell (1998) to introduce a class of dynamic models known as Autoregressive Conditional Duration models, described below.

2.4.1 Autoregressive Conditional Duration (ACD) Models

Let \( \{T_i, i \geq 0\} \) be a sequence of times of occurrence of certain financial events and we assume that \( 0 = T_0 < T_1 < T_2 < \ldots \). Then the \( i^{th} \) duration \( X_i \), the interval between the \( i-1^{th} \) and \( i^{th} \) occurrence of the event is defined by

\[
X_i = T_i - T_{i-1}, \quad i = 1, 2, \ldots
\]

The basic ACD model of Engle and Russell (1998) expresses the seasonally adjusted duration \( X_i \) in the form of a multiplicative error model as

\[
X_i = \psi_i \varepsilon_i, \quad (2.6)
\]

where \( \{\varepsilon_i\} \) is a sequence of independent and identically distributed non-negative random variables with unit mean. Here \( \psi_i \) is introduced as the conditional expectation of the duration \( X_i \) given the information on the past durations. That is, \( \psi_i = E(X_i|F_{i-1}) \), where \( F_{i-1} = \sigma(X_1, X_2, \ldots, X_{i-1}) \) is the sigma field generated by \( (X_1, X_2, \ldots, X_{i-1}) \). Engle and Russell (1998) defined the ACD \((p, q)\) model specifying \( \psi_i \) as

\[
\psi_i = \omega + \sum_{j=1}^{p} \alpha_j X_{i-j} + \sum_{j=1}^{q} \beta_j \psi_{i-j}, \quad (2.7)
\]

where \( p \) and \( q \) are non-negative integers. The following conditions are imposed on the parameters for the stationarity of the sequence:
Different classes of ACD models can be defined either by the choice of the functional form of \( \psi_i \) or by the choice of the distribution for \( \epsilon_i \). For example, ACD models with several standard distributions such as Exponential and Weibull (Engle and Russell (1998)), Burr (Grammig and Maurer (2000)), etc. have been studied in the literature, see Pacurar (2008) for a detailed survey. The equations (2.6) and (2.7) define the dynamical structure of the standard ACD models, which can be viewed as an observation driven model like GARCH model described in Section 2.3. However, in the latter case the model was specified for the conditional variance of the returns.

The standard ACD model has been extended in several ways, directed mainly to improving the fitting of the stylized facts of financial durations. The strong similarity between the ACD and GARCH models nurtured the rapid expansion of alternative specifications of conditional durations. An equally important model for analysing financial durations is a class of parameter driven model introduced by Bauwens and Veredas (2004) known as Stochastic Conditional Duration model.

### 2.4.2 Stochastic Conditional Duration (SCD) Models

In this Section, we analyse a class of parametric models for durations, which are referred to as SCD models proposed by Bauwens and Veredas (2004). In contrast to the ACD model of Engle and Russell (1998), in which the conditional mean of the
duration is modelled as a conditionally deterministic function of past information, the SCD model treats the conditional mean of durations as a stochastic latent process, with innovations to the process captured by an appropriate distribution with positive support. As such, the contrast between the two specifications mimics the contrast between the GARCH and SV frameworks for capturing the conditional volatility of financial returns. In particular, as is the case with the SV model, the SCD model presents a potentially more complex estimation problem than its alternative, by augmenting the set of unknowns with a set of unobservable latent factors.

An SCD model of order one is defined by

\[ X_i = e^{\psi_i} \varepsilon_i, \psi_i = \omega + \beta \psi_{i-1} + u_i, \quad i = 1, 2, ..., \]  

(2.9)

where \( u_i \) follows independent and identically distributed \( N(0, \sigma^2) \) so that \( \{\psi_i\} \) defines a Gaussian AR(1) sequence and \( \varepsilon_i \) is as defined in the case of (2.6). Unlike in the case of ACD models, the analysis of SCD model is more complicated due to the presence of latent variables, \( \psi_i \), which are not observable. Economically, the latent factor can be interpreted as information flow that cannot be observed directly but drives the duration process. In this sense, the SCD model is the counterpart of the SV model introduced by Taylor (1986).

A difficulty associated with SCD framework is the parameter estimation because no explicit expression for the likelihood function of SCD model is directly available due to the presence of latent structure in \( \psi_i \). The evaluation of the likelihood function of the SCD model requires computing an integral that has the dimension of the sample
size. Bauwens and Galli (2009) developed ML estimation based on the efficient importance sampling (EIS) method for computing such integral. Other methods, that are less demanding in computing time, do not evaluate the exact likelihood function. The easiest two techniques are quasi-maximum likelihood (QML) and generalized method of moments (GMM). These techniques provide asymptotically consistent estimators and previous research seems to indicate that the behaviour of the QML estimator is better than the one of GMM in the context of the stochastic volatility model; see Ruiz (1994) and Jacquier et al. (1994). Bauwens and Veredas (2004) used QML based on the transformation of the model into a linear state space representation and the application of the Kalman filter.