1 Introduction

The aim of this book is to provide the researcher in financial markets with the techniques necessary to undertake the empirical analysis of financial time series. To accomplish this aim we introduce and develop both univariate modelling techniques and multivariate methods, including those regression techniques for time series that seem to be particularly relevant to the finance area.

Why do we concentrate exclusively on time series techniques when, for example, cross-sectional modelling plays an important role in empirical investigations of the capital asset pricing model (CAPM; see, as an early and influential example, Fama and MacBeth, 1973)? Moreover, why do we not address the many issues involved in modelling financial time series in continuous time and the spectral domain, although these approaches have become very popular, for example, in the context of derivative asset pricing? Our answer is that, apart from the usual considerations of personal expertise and interest plus constraints on manuscript length, it is because time series analysis, in both its theoretical and empirical aspects, has been for many years an integral part of the study of financial markets.

The first attempts to study the behaviour of financial time series were undertaken by financial professionals and journalists rather than by academics. Indeed, this seems to have become a long-standing tradition, as, even today, much empirical research and development still originates from the financial industry itself. This can be explained by the practical nature of the problems, the need for specialised data and the potential gains from such analysis. The earliest and best-known example of published research on financial time series is by the legendary Charles Dow, as expressed in his editorials in the *Wall Street Times* between 1900 and 1902. These writings formed the basis of 'Dow theory' and influenced what later became known as technical analysis and chartism. Although Dow did not collect and publish his editorials separately, this was done posthumously by his follower Samuel Nelson (Nelson, 1902). Dow's original ideas were later interpreted and further extended by Hamilton (1922) and Rhea (1932). These ideas enjoyed some recognition amongst academics at the time: for example, Hamilton was elected a fellow of the Royal Statistical Society. As characteristically treated by Malkiel (2003), however, technical analysis and chartist approaches became anathema to academics, despite their widespread popularity amongst financial professionals. Although Dow and his followers discussed many of the ideas we encounter in modern finance and time series analysis, including stationarity, market efficiency, correlation between asset returns and indices, diversification and unpredictability, they made no serious effort to adopt formal statistical methods. Most of the empirical analysis involved the painstaking interpretation of detailed charts of sectoral stock price averages, thus forming the celebrated Dow-Jones indices. It was argued that these indices discount all necessary information and provide the best predictor of future events. A fundamental idea, very relevant to the theory of cycles by Stanley Jevons and the 'Harvard A-B-C curve' methodology of trend decomposition by Warren Persons, was that market price variations consisted of three primary movements: daily, medium-term and long-term (see Samuelson, 1987). Although criticism of Dow theory and technical analysis has been a favourite pastime of academics for many years, evidence regarding its merit remains controversial (see, for example, Brown, Goetzmann and Kumar, 1998).

The earliest empirical research using formal statistical methods can be traced back to the papers by Working (1934), Cowles (1933, 1944) and Cowles and Jones (1937). Working focused attention on a previously noted characteristic of commodity and stock prices: namely, that they resemble cumulations of purely random changes. Alfred Cowles 3rd, a quantitatively trained financial analyst and founder of the Econometric Society and the Cowles Foundation, investigated the ability of market analysts and financial services to predict future price changes, finding that there was little evidence that they could. Cowles and Jones reported evidence of positive correlation between successive price changes, but, as Cowles (1960) was later to remark, this was probably due to their taking monthly averages of daily or weekly prices before computing changes: a 'spurious correlation' phenomenon, analysed by Working (1960).

The predictability of price changes has since become a major theme of financial research but, surprisingly, little more was published until Kendall's (1953) study, in which he found that the weekly changes in a wide variety of financial prices could not be predicted either from past changes in the series or from past changes in other price series. This seems to have been the first explicit reporting of this oft-quoted property of financial prices, although further impetus to research on price predictability was provided only by the publication of the papers by Roberts (1959) and Osborne (1959). The former presents a largely heuristic argument as to why successive price changes should be independent, while the latter develops the proposition that it is not absolute price changes but the logarithmic price changes that are independent of each other. With the auxiliary assumption that the changes themselves are normally distributed, this implies that prices are generated as Brownian motion.

The stimulation provided by these papers was such that numerous articles appeared over the next few years investigating the hypothesis that price changes (or logarithmic price changes) are independent, a hypothesis that came to be termed the 'random walk' model, in recognition of the similarity of the evolution of a price series to the random stagger of a drunk. Indeed, the term 'random walk' is believed to have first been used in an exchange of correspondence appearing in *Nature* in 1905 (see Pearson and Rayleigh, 1905), which was concerned with the optimal search strategy for finding a drunk who had been left in the middle of a field at the dead of night! The solution is to start exactly where the drunk had been placed, as that point is an unbiased estimate of the drunk's future position since he will presumably stagger along in an unpredictable and random fashion.

The most natural way to state formally the random walk model is as

$$P_t = P_{t-1} + a_t \tag{1.1}$$

where P_t is the price observed at the beginning of time t and a_t is an error term which has zero mean and whose values are independent of each other. The price change, $\Delta P_t = P_t - P_{t-1}$, is thus simply a_t and hence is independent of past price changes. Note that, by successive backward substitution in (1.1), we can write the current price as the cumulation of all past errors, i.e.

$$P_t = \sum\nolimits_{i=1}^{t} a_i$$

so that the random walk model implies that prices are indeed generated by Working's 'cumulation of purely random changes'. Osborne's model of Brownian motion implies that equation (1.1) holds for the logarithms of P_t and, further, that a_t is drawn from a zero mean normal distribution having constant variance.

Most of the early papers in this area are contained in the collection of Cootner (1964), while Granger and Morgenstern (1970) provide a detailed development and empirical examination of the random walk model and various of its refinements. Amazingly, much of this work had been anticipated by the French mathematician Louis Bachelier (1900; English translation in Cootner, 1964) in a remarkable PhD thesis in which he developed an elaborate mathematical theory of speculative prices, which he then tested on the pricing of French government bonds, finding that such prices were consistent with the random walk model. What made the thesis even more remarkable was that it also developed many of the mathematical properties of Brownian motion that had been thought to have first been derived some years later in the physical sciences, particularly by Einstein! Yet, as Mandelbrot (1989) remarks, Bachelier had great difficulty in even getting himself a university appointment, let alone getting his theories disseminated throughout the academic community! The importance and influence of Bachelier's path-breaking work is discussed in Sullivan and Weithers (1991) and Dimand (1993).

It should be emphasised that the random walk model is only a hypothesis about how financial prices move. One way in which it can be tested is by examining the autocorrelation properties of price changes: see, for example, Fama (1965). A more general perspective is to view (1.1) as a particular model within the class of autoregressive integrated moving average (ARIMA) models popularised by Box and Jenkins (1976). Chapter 2 thus develops the theory of such models within the general context of (univariate) linear stochastic processes. An important aspect of specifying ARIMA models is to be able to determine correctly the order of integration of the series being analysed and, associated with this, the appropriate way of modelling trends and structural breaks. To do this formally requires the application of unit root tests and a vast range of related procedures. Tests for unit roots and alternative trend specifications are the focus of chapter 3.

We should avoid giving the impression that the only financial time series of interest are stock prices. There are financial markets other than those for stocks, most notably for bonds and foreign currency, but there also exist the various futures, commodity and derivative markets, all of which provide interesting and important series to analyse. For certain of these, it is by no means implausible that models other than the random walk may be appropriate, or, indeed, models from a class other than the ARIMA. Chapter 4 therefore discusses various topics in the general analysis of linear stochastic models: for example, methods of decomposing an observed series into two or more unobserved components and of determining the extent of the 'memory' of a series, by which is meant the behaviour of the series at low frequencies or, equivalently, in the very long run. A variety of examples taken from the financial literature are provided throughout these chapters.

The random walk model has been the workhorse of empirical finance for many years, mainly because of its simplicity and mathematical tractability. Its prominent role was also supported by theoretical models that obtained unpredictability as a direct implication of market efficiency, or, more broadly speaking, of the condition whereby market prices fully, correctly and instantaneously reflect all the available information. An evolving discussion of this research can be found in a series of papers by Fama (1970, 1991, 1998), while Timmermann and Granger (2004) address market efficiency from a forecasting perspective. As LeRoy (1989) discusses, it was later shown that the random walk behaviour of financial prices is neither a sufficient nor a necessary condition for rationally determined financial prices. Moreover, the assumption in (1.1) that price changes are independent was found to be too restrictive to be generated within a reasonably broad class of optimising models. A model that is appropriate, however, can be derived for stock prices in the following way (similar models can be derived for other sorts of financial prices, although the justification is sometimes different: see LeRoy, 1982). The return on a stock from t to $t+1$ is defined as the sum of the dividend yield and the capital gain – i.e. as

$$r_{t+1} = \frac{P_{t+1} + D_t - P_t}{P_t} \tag{1.2}$$

where D_t is the dividend paid during period t. Let us suppose that the expected return is constant, $E_t(r_{t+1}) = r$, where $E_t(\ )$ is the expectation conditional on information available at t: r_t is then said to be a *fair game*. Taking expectations at t of both sides of (1.2) and rearranging yields

$$P_t = (1+r)^{-1} E_t(P_{t+1} + D_t) \tag{1.3}$$

which says that the stock price at the beginning of period t equals the sum of the expected future price and dividend, discounted back at the rate r. Now assume that there is a mutual fund that holds the stock in question and that it reinvests dividends in future share purchases. Suppose that it holds h_t shares at the beginning of period t, so that the value of the fund is $x_t = h_t P_t$. The assumption that the fund ploughs back its dividend income implies that h_{t+1} satisfies

$$h_{t+1} P_{t+1} = h_t(P_{t+1} + D_t)$$

Thus

$$E_t(x_{t+1}) = E_t(h_{t+1} P_{t+1}) = h_t E_t(P_{t+1} + D_t) = (1+r) h_t P_t = (1+r) x_t$$

i.e. x_t is a *martingale* (if, as is common, $r > 0$, we have $E_t(x_{t+1}) \geq x_t$, so that x_t is a *submartingale*; LeRoy (1989, pp. 1593–4) offers an example, however, in which r could be negative, in which case x_t will be a *supermartingale*). LeRoy (1989) emphasises that price itself, without dividends added in, is not generally a martingale, since from (1.3) we have

$$r = E_t(D_t)/P_t + E_t(P_{t+1})/P_t - 1$$

so that only if the expected dividend/price ratio (or dividend yield) is constant, say $E_t(D_t)/P_t = d$, can we write P_t as the submartingale (assuming $r > d$)

$$E_t(P_{t+1}) = (1 + r - d)P_t$$

The assumption that a stochastic process – y_t, say – follows a random walk is more restrictive than the requirement that y_t follows a martingale. The martingale rules out any dependence of the conditional expectation of Δy_{t+1} on the information available at t, whereas the random walk rules out not only this but also dependence involving the higher conditional moments of Δy_{t+1}. The importance of this distinction is thus evident: financial series are known to go through protracted quiet periods and also protracted periods of turbulence. This type of behaviour could be modelled by a process in which successive conditional variances of Δy_{t+1} (but *not* successive levels) are positively autocorrelated. Such a specification would be consistent with a martingale, but not with the more restrictive random walk.

Martingale processes are discussed in chapter 5, and lead naturally on to non-linear stochastic processes that are capable of modelling higher conditional moments, such as the autoregressive conditionally heteroskedastic (ARCH) model introduced by Engle (1982) and stochastic variance models. Related to these models is the whole question of how to model volatility itself, which is of fundamental concern to financial modellers and is therefore also analysed in this chapter. Of course, once we entertain the possibility of non-linear generating processes a vast range of possible processes become available, and those that have found, at least potential, use in modelling financial time series are developed in chapter 6. These include bilinear models, Markov switching processes, smooth transitions and chaotic models. The chapter also includes a discussion of computer intensive techniques such as non-parametric modelling and artificial neural networks. An important aspect of nonlinear modelling is to be able to test for nonlinear behaviour, and testing procedures thus provide a key section of this chapter.

The focus of chapter 7 is on the unconditional distributions of asset returns. The most noticeable future of such distributions is their leptokurtic property: they have fat tails and high peakedness compared to a normal distribution. Although ARCH processes can model such features, much attention in the finance literature since Mandelbrot's (1963a, 1963b) path-breaking papers has concentrated on the possibility that returns are generated by a stable process, which has the property of having an infinite variance. Recent developments in statistical analysis have allowed a much deeper investigation of the tail shapes of empirical distributions, and methods of estimating tail shape indices are introduced and applied to a variety of returns series. The chapter then looks at the implications of fat-tailed distributions for testing the covariance stationarity assumption of time series analysis, data analytic methods of modelling skewness and kurtosis, and the impact of analysing transformations of returns rather than the returns themselves.

The remaining three chapters focus on multivariate techniques of time series analysis, including regression methods. Chapter 8 concentrates on analysing the relationships between a set of *stationary* – or, more precisely, *non-integrated* – financial time series and considers such topics as general dynamic regression, robust estimation, generalised methods of moments, multivariate regression, ARCH-in-mean and multivariate ARCH models, vector autoregressions, Granger causality, variance decompositions and impulse response analysis. These topics are illustrated with a variety of examples drawn from the finance literature: using forward exchange rates as optimal predictors of future spot rates; modelling the volatility of stock returns and the risk premium in the foreign exchange market; testing the CAPM; and investigating the interaction of the equity and gilt markets in the United Kingdom.

Chapter 9 concentrates on the modelling of *integrated* financial time series, beginning with a discussion of the spurious regression problem, introducing cointegrated processes and demonstrating how to test for cointegration, and then moving on to consider how such processes can be estimated. Vector error correction models are analysed in detail, along with associated issues in causality testing and impulse response analysis, alternative approaches to testing for the presence of a long-run relationship, and the analysis of both common cycles and trends. The techniques introduced in this chapter are illustrated with extended examples analysing the market model and the interactions of the UK financial markets.

Finally, chapter 10 considers modelling issues explicit to finance. Samuelson (1965, 1973) and Mandelbrot (1966) have analysed the implications of equation (1.3), that the stock price at the beginning of time t

equals the discounted sum of the next period's expected future price and dividend, to show that this stock price equals the expected discounted, or present, value of *all* future dividends – i.e. that

$$P_t = \sum\nolimits_{i=0}^{\infty} (1+r)^{-(i+1)} E_t(D_{t+i}) \tag{1.4}$$

which is obtained by recursively solving (1.3) forwards and assuming that $(1+r)^{-n}E_t(P_{t+n})$ converges to zero as $n \to \infty$. Present value models of the type (1.4) are analysed comprehensively in chapter 10, with the theme of whether stock markets are excessively volatile, perhaps containing speculative bubbles, being used extensively throughout the discussion and in a succession of examples, although the testing of the expectations hypothesis of the term structure of interest rates is also used as an example of the general present value framework. The chapter also discusses recent research on non-linear generalisations of cointegration and how structural breaks may be dealt with in cointegrating relationships.

Having emphasised earlier in this chapter that the book is exclusively about modelling financial time series, we should state at this juncture what the book is not about. It is certainly not a text on financial market theory, and any such theory is discussed only when it is necessary as a motivation for a particular technique or example. There are numerous texts on the theory of finance, and the reader is referred to these for the requisite financial theory: two notable texts that contain both theory and empirical techniques are Campbell, Lo and MacKinlay (1997) and Cuthbertson (1996). Neither is it a textbook on econometrics. We assume that the reader already has a working knowledge of probability, statistics and econometric theory, in particular least squares estimation. Nevertheless, it is also non-rigorous, being at a level roughly similar to Mills (1990), in which references to the formal treatment of the theory of time series are provided.

When the data used in the examples throughout the book have already been published, references are given. Previous unpublished data are defined in the data appendix, which contains details on how they may be accessed. All standard regression computations were carried out using *EVIEWS 5.0* (EViews, 2003), but use was also made of *STAMP 5.0* (Koopman *et al.*, 2006), *TSM 4.18* (Davidson, 2006a) and occasionally other econometric packages. 'Non-standard' computations were made using algorithms written by the authors in *GAUSS* and *MatLab*.

2 Univariate linear stochastic models: basic concepts

Chapter 1 has emphasised the standard representation of a financial time series as that of a (univariate) linear stochastic process, specifically as being a member of the class of ARIMA models popularised by Box and Jenkins (1976). This chapter provides the basic theory of such models within the general framework of the analysis of linear stochastic processes. As already stated in chapter 1, our treatment is purposely non-rigorous. For detailed theoretical treatments, but which do not, however, focus on the analysis of financial series, see, for example, Brockwell and Davis (1996), Hamilton (1994), Fuller (1996) or Taniguchi and Kakizawa (2000).

2.1 Stochastic processes, ergodicity and stationarity

2.1.1 Stochastic processes, realisations and ergodicity

When we wish to analyse a financial time series using formal statistical methods, it is useful to regard the observed series, $(x_1, x_2, \ldots, x_T)$, as a particular *realisation* of a stochastic process. This realisation is often denoted $\{x_t\}_1^T$, while, in general, the stochastic process itself will be the family of random variables $\{X_t\}_{-\infty}^{\infty}$ defined on an appropriate probability space. For our purposes it will usually be sufficient to restrict the index set $\mathrm{T} = (-\infty, \infty)$ of the parent stochastic process to be the same as that of the realisation, i.e. $\mathrm{T} = (1, T)$, and also to use x_t to denote both the stochastic process and the realisation when there is no possibility of confusion.

With these conventions, the stochastic process can be described by a T-dimensional probability distribution, so that the relationship between a realisation and a stochastic process is analogous to that between the sample and population in classical statistics. Specifying the complete form of the probability distribution will generally be too ambitious a task, and we usually content ourselves with concentrating attention on the first and second moments: the T means

$$E(x_1), E(x_2), \ldots, E(x_T)$$

T variances

$$V(x_1), V(x_2), \ldots, V(x_T)$$

and $T(T-1)/2$ covariances

$$Cov\left(x_i, x_j\right), \quad i < j$$

If we could assume joint normality of the distribution, this set of expectations would then completely characterise the properties of the stochastic process. As we shall see, however, such an assumption is unlikely to be appropriate for many financial series. If normality cannot be assumed but the process is taken to be *linear*, in the sense that the current value of the process is generated by a linear combination of previous values of the process itself and current and past values of any other related processes, then, again, this set of expectations would capture its major properties. In either case, however, it will be impossible to infer all the values of the first and second moments from just one realisation of the process, since there are only T observations but $T+T(T+1)/2$ unknown parameters. Hence, further simplifying assumptions must be made to reduce the number of unknown parameters to more manageable proportions.

We should emphasise that the procedure of using a single realisation to infer the unknown parameters of a joint probability distribution is valid only if the process is *ergodic*, which essentially means that the sample moments for finite stretches of the realisation approach their population counterparts as the length of the realisation becomes infinite. For more on ergodicity, see, for example, Granger and Newbold (1986, chap. 1) or Hamilton (1994, chap. 3.2) and, since it is difficult to test for ergodicity using just (part of) a single realisation, it will be assumed from now on that all time series have this property. Domowitz and El-Gamal (2001) have provided a set of sufficient assumptions under which a single time series trajectory will contain enough information to construct a consistent non-parametric test of ergodicity.

2.1.2 Stationarity

One important simplifying assumption is that of *stationarity*, which requires the process to be in a particular state of 'statistical equilibrium' (Box and