

# ON MODELLING SPECULATIVE PRICES: THE EMPIRICAL LITERATURE

Elena Andreou

*University of Manchester*

Nikitas Pittis

*University of Piraeus*

Aris Spanos

*Virginia Tech*

**Abstract.** Traditionally, financial theory and in particular asset pricing models have assumed (implicitly or explicitly) a certain probabilistic structure for speculative prices. The probabilistic structure is usually defined in terms of specific statistical models and relates to the dependence, heterogeneity and the distribution of such prices. The primary objective of this paper is to trace the development of various statistical models proposed since Bachelier (1900), in an attempt to assess how well these models capture the empirical regularities exhibited by data on speculative prices.

**Keywords.** Stock market returns; Volatility; Efficient market hypothesis; ARCH

## 1. Introduction

Financial theories are often based on explicit or implicit assumptions concerning the probabilistic structure of the stochastic processes underlying the speculative price data (e.g. stock returns, exchange rates, interest rates etc.). For instance, the notion of the efficient market hypothesis presupposes that speculative prices can be modelled as *random walks* (Bachelier, 1900; Fama, 1970) or *martingale processes* (Mandelbrot, 1967; Samuelson, 1965). Moreover, theoretical asset pricing models, such as:

- (i) the Mean-Variance Theorem (Markowitz, 1952),
- (ii) the Capital Asset Pricing Model (CAPM) (Sharpe, 1964; Lintner, 1965, Mossin, 1966; Merton, 1973),
- (iii) the Consumption CAPM (Lucas, 1978; Breeden, 1979),
- (iv) the non-expected utility CAPM (Epstein and Zin, 1991), and
- (v) the Black-Scholes options pricing model (Black and Scholes, 1973),

are based on probabilistic assumptions such as *Normality*, *uncorrelatedness* and *stationarity*, relating to the returns (log-price changes). Despite that, the empirical evidence accumulated over the last few decades indicates departures from these probabilistic assumptions. Data on speculative prices often exhibit *bell shape symmetry*, *leptokurticity*, *non-linear temporal dependence* and *erratic behaviour* of sample second moments (e.g. Kendall, 1953; Mandelbrot, 1963; for a recent summary see Bollerslev *et al.*, 1994 and Pagan, 1996).

The primary objective of the paper is to discuss the various empirical models of speculative price data since Bachelier (1900), with particular attention to how well the various statistical models were able to capture the empirical regularities as perceived at the time. We note at the outset that a statistical model in this context is considered to be simply a consistent set of probabilistic assumptions. In order to present a coherent account of the empirical literature we adopt the basic taxonomy of probabilistic assumptions into

- [1] Distribution, [2] Dependence, [3] Heterogeneity. (1)

We examine how these assumptions relate to the empirical regularities of asset returns (see Spanos, 1986, 1999) and how they assist in specifying the relevant models for stock returns. It is shown below that several empirical regularities either went largely unnoticed or were misinterpreted.

In relation to the modelling of speculative prices, one can identify three distinct periods since the early twentieth century. The Bachelier-Kendall (1900–1960) era commenced in 1900 with the Brownian motion (Bachelier, 1900, 1914) and continued with the Normal random walk model. The stylized facts of ‘approximate’ *Normality* and *uncorrelatedness* were found in certain data on returns (see for instance Kendall, 1953). The Mandelbrot (1960–1980) era revisited the stylized facts of the previous period and brought out three interrelated inconsistencies within the Normal Random Walk model: *leptokurticity*, *erratic behaviour of the sample variance* and *volatility clustering*. In an attempt to deal with the first two inconsistencies, the Normality assumption was replaced by that of the Pareto-Levy family of distributions. Finally, the dynamic volatility era commenced in the early eighties with the introduction of the Autoregressive Conditional Heteroscedasticity models (Engle, 1982) and concentrated mostly on modelling volatility dynamics. In this respect, a number of alternative functional forms of the conditional variance were proposed, thus giving rise to a new family of time series models (surveyed in Bollerslev *et al.*, 1992, 1994).

## 2. The Bachelier-Kendall era (1900–1960)

A collection of some of the early classic and influential papers in financial market analysis is presented in the book edited by Cootner (1964). The first period of research in speculative prices (1900–1960) is called after the French mathematician Bachelier (1900, 1914), who contributed the theoretical random walk model for the analysis of speculative prices. The empirical modelling of speculative prices

data begins in the mid thirties with a study by Working (1934) which concludes that stock returns behave like ‘*numbers from a lottery*’. The *unpredictability* of stock returns was taken up again in the 1950s (see for instance, Kendall, 1953; Roberts, 1959; Working, 1960; Cowles, 1960; Moore, 1962; Granger and Morgenstern, 1970). Cootner summarizes the general consensus among academics about the unpredictability in stock prices as follows:

‘The time (spent on analysing stock prices) might just as well be spent in analysing the results of a fair roulette wheel’ (ibid. p. 232).

From the theory point of view, the unpredictability of returns (lack of *patterns*) was later justified by the efficient market hypothesis (Fama, 1965,1970, Samuelson, 1965), which utilizes the assumption of instantaneous adjustment of prices to new information. Let  $\{p_t = \ln P_t, t \in \mathbb{T}\}$  and  $\{r_t, t \in \mathbb{T}\}$  denote the stock price and returns processes, respectively, where changes in log-prices,  $\Delta p_t$ , are by definition equal to returns:

$$r_t \equiv \Delta p_t = p_t - p_{t-1}, \quad t \in \mathbb{T}. \tag{2}$$

The early literature (see Kendall, 1953 and Roberts, 1959) views the Random Walk formulation as a system which generates the stock price process via:

$$p_t = p_{t-1} + r_t, \quad r_t \sim \text{IID}(0, \sigma^2), \quad t \in \mathbb{T}, \tag{3}$$

where  $r_t$  is an Independent and Identically Distributed (IID) process with zero mean and constant variance  $\sigma^2$ . In view of the randomness of the sample, this literature often calls (3) the ‘*random sample model*’ or the ‘*chance mechanism*’. This relationship implies that prices can be viewed as *partial sums* of returns in the sense that:

$$p_t = \sum_{k=1}^t r_k, \quad t \in \mathbb{T}. \tag{4}$$

Two things are worth noting at this stage. First, this early literature makes no explicit distributional assumption. Second, by construction  $\{p_t, t \in \mathbb{T}\}$  is a Markov-dependent process whose first two moments exist. In the literature of this period, however, it is often implicitly assumed that the underlying distribution of returns is Normal. Hence, the statistical model is the Normal Random Walk:

$$p_t = p_{t-1} + r_t, \quad r_t \sim \text{NIID}(0, \sigma^2), \quad t \in \mathbb{T}, \tag{5}$$

where N denotes Normality. For this model, the process  $\{p_t, t \in \mathbb{T}\}$  is Markov with a probabilistic structure given by:

$$\begin{pmatrix} p_t \\ p_{t-1} \end{pmatrix} \sim \text{N} \left[ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2 t & \sigma^2(t-1) \\ \sigma^2(t-1) & \sigma^2(t-1) \end{pmatrix} \right], \quad t \in \mathbb{T}. \tag{6}$$

This is the discrete-time equivalent to the Brownian motion process proposed by Bachelier (1900, 1914). Bachelier was the first to provide the law of probability for stock market fluctuations. His analysis begins with the definition of the 'total mathematical expectation' of a player, as the sum of the products of the uncertain gains with the corresponding probabilities of their occurring. Then the game is fair if the total mathematical expectation for the representative player is zero; this is based on Bachelier's intuition that prices follow a martingale. Bachelier's analysis is based on the following assumptions:

- (i) Stock returns follow a fair game: 'the market, at a given instant, considers not only currently negotiable transactions, but even those which will be based on a subsequent fluctuation in prices as having a zero expectation' (Bachelier, 1900; see Cootner (1964), p. 28).
- (ii) The probability that the price  $p_{t+n}$  will be quoted at a given future moment,  $t+n$ , is a function of the current price,  $p_t$ . This assumption introduces a Markov-type dependence into the process of prices. Note that neither the notion of Markovness nor that of a martingale existed at the time.
- (iii) Bachelier also assumes that transactions are fairly uniformly spread across time, the distribution of price changes between successive transactions has finite variance and the number of transactions within a particular time interval is very large. The latter assumption is needed for the utilization of the Central Limit theorem to call upon Normality.

Under these conditions, Bachelier proves the following: First, the conditional probability that price,  $p_{t+n}$ , will be quoted at the moment  $t+n$ , given that price  $p$  has been quoted at the moment  $t$ , is governed by the Gaussian law. Second, the unconditional probability, as a function of time, is also Gaussian and proportional to the square root of time. This result was rediscovered five years later by Einstein in a different context.

The analogy between the temporal behaviour of stock prices and the kinetic molecular theory was also demonstrated by Osborne (1959) in a rigorous way. In particular, Osborne showed that the logarithms of stock prices can be thought of as an ensemble of decisions in statistical equilibrium and that this ensemble of logarithms of prices is analogous to the ensemble of coordinates of a large number of molecules. More specifically, the steady state distribution function of the first differences of the logarithms of prices was shown to be precisely the probability distribution of a particle in Brownian motion. Sufficient but not necessary conditions to derive this distribution require prices to be related to the subjective sensation of values by the Weber-Fechner law and the opportunity to profit to be equal between buyers and sellers.

The Brownian motion theory has gone through several phases in the continuous-time financial models and continues to this day (for a full classification of continuous time — financial models, see Sawyer, 1993). The most important modifications include the following: (i) The Geometric Brownian motion, first proposed by Samuelson (1965), which constitutes the basis for the

Black-Scholes option pricing model, (ii) the Arithmetic Brownian motion, adopted by Vasicek (1977) in his model of the term structure, (iii) the Brownian bridge, used by Ball and Toulos (1983) to price default-free discount bonds, (iv) the Square Root process, put forward by Cox, Ingersoll and Ross (1985) in their study of interest rates, (v) the heteroskedastic process, adopted by Schaefer and Schwartz (1987) in the pricing of bond options and (vi) the stochastic volatility process in which the variance of the process itself is modelled as a Brownian motion (see, for example, Chesney and Scott, 1989), (vii) the Fractional Brownian motion which accounts for long-range dependence usually observed in the actual stock price data.

The Random Walk model (3) for stock prices has been the cornerstone of all empirical analysis in the Bachelier era and the early foundation of the efficient market hypothesis (EMH). In relation to the latter Fama (1965,1970) argues that:

‘In the early treatments of the efficient market model, the statement that the current price of a security fully reflects available information was assumed to imply that price changes or one-period returns are independent. In addition it was usually assumed that returns are identically distributed. Together the two hypotheses constitute the random walk model’, (Fama, 1970, p. 386).

Kendall (1953) related the IID assumptions to the lack of any *pattern* in the sequence of returns as follows:

‘...changes in security prices behaved nearly as if they have been generated by a suitably designed roulette wheel for which each outcome was statistically independent of past history and for which relative frequencies were reasonably stable over time’, (ibid. p. 754).

In the 1950s the empirical literature on speculative prices confronted the Random Walk model (3) with data on speculative price by raising three important inconsistencies:

### 2.1. *Non-Identically Distributed*

The first inconsistency was the departure from the identically distributed assumption due to an apparent changing variance in stock returns data. In particular, Kendall (1953) divided the sample of the Chicago Wheat series into two sub-samples to examine homogeneity over time. Comparison of the variance of the two sub-samples suggested that:

‘there has been an increase in variability since World War I’ and that returns are a ‘rather unusual case of a time-series for which the mean remains constant but the variance appears to be increasing’ implying that the sample mean is  $t$ -invariant whereas the variance is  $t$ -dependent (Kendall, 1953, in Cootner, 1964 p. 89).

This observation led to the rejection of the Random Walk model in favour of the Heterogeneous 'Random Walk' model:

$$p_t = p_{t-1} + r_t, \quad r_t \sim \text{NI}(0, \sigma_t^2), \quad t \in \mathbb{T}, \quad (7)$$

where  $r_t$  is Independent but non-Identically Distributed. Note that the name 'Random Walk' is somewhat misleading because the probabilistic literature retains the term only when the process  $\{r_t, t \in \mathbb{T}\}$  is IID.

The time dependence of the empirical variance was first noted by Kendall (1953) and subsequently by Mandelbrot (1963). More recently, Schwert (1989), Phillips and Loretan (1994) and Campbell *et al.*, (1997) argued that the assumption of ID increments is not plausible for financial asset prices over long time spans. This may be attributed to the countless changes in economic, technological and institutional environment of stock prices. Note that according to the model specified by (7) the variance of returns is time-variant (e.g. see Schwert, 1989), but not heteroskedastic, because the relevant distribution for returns is the marginal and not the conditional. In the context of the Random Walk models, returns is an independent process which does not allow the existence of dynamic conditional moments.

## 2.2. Temporal dependence

The core assumption of the statistical models during this period is the *temporal independence* of the returns process. Tests for independence were carried out in a Gaussian framework, thus making independence equivalent to uncorrelatedness. Indeed, Kendall (1953), Cowles (1960), Moore (1962), Granger and Morgenstern (1970) test the independence assumption focusing on serial correlations.

Kendall (1953) found that stock returns were serially uncorrelated in the case of the Chicago Wheat weekly series. However, he finds that the British Industrial weekly share *index* prices were serially correlated and the New York cotton *monthly* prices seemed to follow a Markov process. The last piece of evidence is dismissed by arguing that:

'In some cases the general picture of independence was presented, but in others there were signs of dependence as reflected in the serial correlations'. Yet, 'such serial correlation as is present in these series is so weak as to dispose at once of any possibility of being able to use them for prediction' (Kendall, 1953, in Cootner, 1964, p. 92).

All these 'stylized facts' of stock returns were also systematically reviewed by Moore (1962) who finds that autocorrelation coefficients were 'uniformly small' and usually 'quite insignificant'. At the same time, Moore also argues that:

'...successive changes of the logs of S&P stock price index evidence slight positive dependence' (see Tables 1 and 2 in Moore, 1962, presented in Cootner, 1964 p. 145–6.)

Similarly, Cowles and Jones (1937) found significant evidence of serial correlation in averaged time series indices of stock prices. The presence of this serial correlation has already been explained by Slutsky (1937) who showed that this averaging may produce '*spurious correlation*'. In fact, this point was also made by Alexander (1961) who showed that Kendall's finding about significant auto-correlation was actually spurious, due to the fact that each monthly observation of the cotton price series was an average of four or five weekly observations of the corresponding month. He also pointed out another factor which may introduce spurious correlations into the price changes, namely the fact that the probability of a rise is not necessarily equal to a half. He suggests that any deviations between the random walk model and the actual behaviour of stock prices are totally spurious:

'It must be concluded that the month to month movement of stock prices, at least in direction, is consistent with the hypothesis of a random walk with about a 6 to 4 probability of a rise. Evidence to the contrary was spurious, arising from the correlations introduced by monthly averaging or neglect of the unequal probability of rise and fall' (Alexander, 1961, p. 214).

Similarly Working (1960) showed that the serial correlation in monthly returns could be explained by the averaging of the series from weekly to monthly observations, and from individual stocks to indices. Therefore, Cowles (1960) revisits the previous results (in Cowles and Jones, 1937) using weekly indices based on Wednesday closing prices for each week and monthly indices of middle-of-the-month or beginning-of-the-month prices. Using the sequences and reversals test he *still* finds mixed temporal dependence results. Cootner's conclusion is typical of this blurred picture:

'...some of the evidence supported randomness of stock price changes, while other evidence indicated correlations which suggested that the market was less than perfect' (Cootner, 1964, p. 80–81).

Cootner (1962) also presents some tentative evidence against the independence hypothesis by showing that a specific decision rule does substantially better than random buying of stocks. This rule is based on the assumption of the presence of two disparate groups of market participants. The first group is characterized by a low degree of knowledgeable ability of the market, whereas the second group consists of professional investors. This asymmetry in the possession of information accompanied with some rigidity in the expectations of professionals give rise to the 'random walk with reflecting barriers' model for stock prices. This model implies the presence of tendencies (trends) near the maximum and minimum for price changes to move toward the mean. In addition, for such a process the kurtosis coefficient should be greater than 3. Cootner also shows that the kurtosis coefficient will approach that of a rectangular distribution (a platykurtic distribution) in the limit, if a single trend is involved. These results, however, are subject to the caveat that are purely descriptive, since there is no indication for their statistical significance. This limitation was dealt with by Steiger (1964) who

builds on Cootner's idea that stock prices follow a constrained diffusion process. He compares the actual prices with those produced by a simulated random walk with reflecting barriers and provides a formal test for 'randomness within barriers' which is based on the Kolmogorov-Smirnov test for a common distribution. He finds that the actual and simulated distributions are significantly different at the 2.5% level. It must be noted, however, that the presence of 'sophisticated traders' in the market is sometimes used in favour of the random walk hypothesis. In particular, their actions are assumed to neutralize the dependence in the noise-generating process, which might arise because of self-fulfilling prophecies about the intrinsic value of the stock from the uninformed group (see Fama, 1965).

In an important paper, Osborne (1962) reexamines the general topic of deviations of stock prices from a simple random walk along several new directions. The most important result from his analysis is based on the rejection of the null hypothesis of uniform volume of transactions over time. In fact, it is shown that some intervals of time have a much greater probability to contain transactions than others of equal length. Osborne interpreted this evidence as '*a tendency for stock to be traded in concentrated bursts*' (p. 273). This evidence is consistent with Mandelbrot (1963) although the interpretation is quite different. In particular, Osborne interprets these bursts of occurrences as evidence against the random walk hypothesis, whereas Mandelbrot maintains the independence assumption and puts forward the idea of independent returns which are distributed according to the Stable family of distributions (to be discussed in the next section).

The issue of the presence or absence of correlation in speculative price data remained unresolved until the early 1980s when the presence of higher order temporal dependence was also detected. A retrospective criticism of the temporal dependence testing strategy during this period, is that dependence and correlation were viewed as equivalent, ignoring the possibility of *higher-order temporal dependence*. The latter possibility, however, presupposes departure from the Normality assumption, and movement towards a non-standard regression-based statistical territory. It was not until the beginning of the eighties that such a task was seriously pursued. Nevertheless, Alexander (1961) implicitly acknowledges that non-correlation does not necessarily imply independence and tests the null hypothesis of 'randomness' using a *runs* test. Being unable to reject the null, he concluded that the random walk model best fits the data. Retrospectively, we can interpret the various attempts to rationalize the presence or absence of serial correlation as ex post theory rationalisations of nebulous statistical evidence.

### 2.3. *Leptokurticity*

The third inconsistency concerned the distribution of returns. Kendall (1953) found that the bivariate frequency distribution of weekly price changes (of the Chicago Wheat series) is described by:



'nearly perfect symmetry and an appearance of approximate normality' since 'the distributions are accordingly rather leptokurtic...' (Kendall, 1953, in Cootner, 1964, p. 87).

As confirming evidence he used the sample skewness and kurtosis coefficients:  $\hat{\alpha}_3 = 0.219$ ,  $\hat{\alpha}_4 = 8.506$ ; compared with  $\alpha_3 = 0$ ,  $\alpha_4 = 3$  in the case of the Normal distribution. This apparent departure from the Normal distribution was largely ignored by the early literature, due to a misinterpretation of the Central Limit Theorem (CLT) effect. For instance, Kendall (1953) finds that the frequency distribution of the *monthly* average of the Chicago wheat series, was also symmetric and leptokurtic (given  $\hat{\alpha}_3 = 0.002$  and  $\hat{\alpha}_4 = 5.701$ , p. 91). Nevertheless, he states:

'The distribution is nearer to normality as one might expect under the central limit effect'. Therefore, he infers that 'there is nothing to be gained by taking averages of the monthly figures to obtain an annual figure. Under the Central Limit effect the resulting series would be nearly normal; and successive values would almost certainly be independent', (Kendall, 1953, in Cootner, 1964, p. 91–2).

Calling upon the CLT to justify the use of Normality when going from weekly to monthly and annual data is inappropriate because the *summation* involves only a finite number of terms which *does not* go to infinity (see Spanos, 1999). Also note that this averaging is inappropriate because, as pointed above, it creates spurious correlation (Slutsky, 1937; Working, 1960). Indeed, during the latter part of the Bachelier-Kendall period emphasis was placed on the question of presence or absence of temporal dependence.

In conclusion, the Bachelier period ended leaving largely unresolved three empirical issues concerning the stock returns process:

- (i) Is the distribution Normal?
- (ii) Is the temporal independence assumption valid?
- (iii) Is the identically distributed assumption justified?

The available evidence pointed towards the inappropriateness of the Normal Random Walk model. The presence of leptokurticity calls into question the Normality assumption. The 'blurred' picture relating to the presence of temporal correlation, when viewed in conjunction with non-Normality, calls into question the assumption of temporal independence. The apparent non-stationarity of the variance of returns implies that the assumption of identically distributed returns also seems inappropriate.

### 3. The Mandelbrot era (1960–1980): stable paretian and martingale models

In the 1960s, Mandelbrot published several influential papers on speculative prices shifting the interest away from temporal dependence to the *distributional* features of such data series. Mandelbrot (1963) showed that Bachelier's Brownian

motion model was incompatible with the 'revised' stylized facts on speculative prices.

### 3.1. *Non-Normality*

First, the histograms of returns showed bell-shape symmetry but more peakedness relative to the Gaussian distribution:

'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 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.', (Mandelbrot, 1963, p. 394).

An empirical illustration of the leptokurtic distribution exhibited by stock return data is provided. We use daily data on the U.S. Standards and Poors 500 index (SP500) and the U.K. Financial Times All Share (FT-ALL) stock index. The sample period is 8/1/1988 — 30/5/1997 and the sample size is  $T = 2348$  observations. (Datastream is the data source.) The return on the market portfolios was taken as the log differences of the respective daily stock price index (excluding dividends).

The distributional characteristics of the FT-ALL and SP500 returns series are examined. (Note that the particular day effects encountered in stock returns data have been dummied out.) The descriptive statistics reported in Table 1 do not support the marginal Normality hypothesis on the basis of excess kurtosis. Note that the FT-ALL is also skewed (an empirical regularity also addressed in the recent literature, see, for instance, Mills, 1995). Similarly, the D'Agostino and Stephen (1986) skewness-kurtosis Normality test does not provide support for the Normality assumption in the SP500 and FT-ALL series due to excess kurtosis. In addition, the univariate nonparametric kernel estimates (e.g. Silverman, 1986) of

**Table 1.** Descriptive Statistics for the US & UK Stock Market Returns

	Mean	Standard Deviation	Skewness	Kurtosis	D'Agostino Skewness	D'Agostino Kurtosis
SP500	0.0005	0.007	-0.074	5.255	-1.299 (0.097)	10.895 (0.000)**
FT-ALL	0.0003	0.0068	-0.100	4.465	-2.153 (0.016)**	8.684 (0.000)**

Notes:

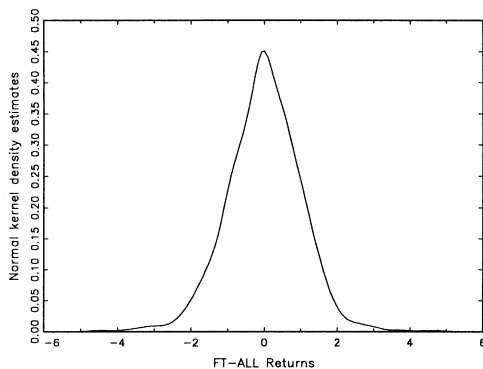
1. The mean, standard deviation, skewness & kurtosis refer to sample statistics.
2. The D'Agostino skewness kurtosis Normality test is performed and the numbers in brackets refer to  $p$ -values.
3. (\*\*) refer to rejection of the null hypothesis at the 5% significance level.

the stock return indices shown in Figures 1 and 2 deviate from the standard Normal,  $N(0, 1)$ , distribution since they appear relatively more peaked. The kernels are symmetric and the sample skewness coefficients are close to zero.

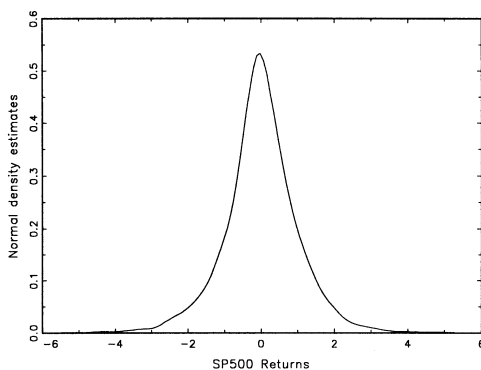
3.2. *Infinite variance and volatility clustering*

The second stylized fact observed by Mandelbrot (1963,1969) refers to the infinite variance syndrome of stock returns which was called the Noah effect (Mandelbrot,1969) suggesting that:

‘...the tails of the distributions of price changes are in fact so extraordinarily long that the sample second moments typically vary in an erratic fashion’, (Mandelbrot, 1963, p. 395).



**Figure 1.** Kernel for FT-ALL.



**Figure 2.** Kernel for SP500.

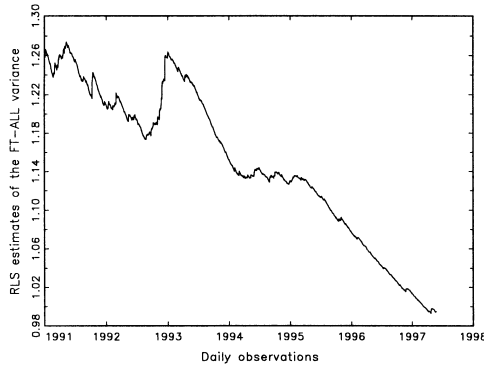
This is examined using the sample recursive variance:

$$\frac{1}{k} \sum_{i=1}^k (r_i - \bar{r})^2, \quad k = 2, 3, \dots, T.$$

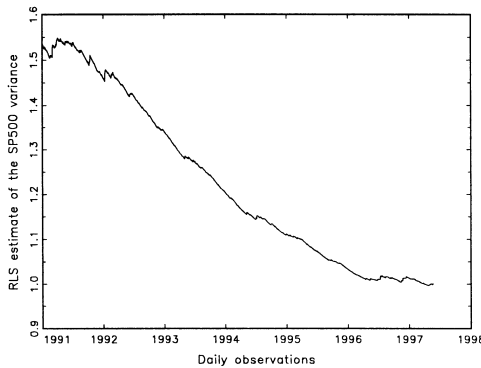
On the basis of the sample recursive variance behaviour, Mandelbrot (1963) advocates that:

‘...the second moment ... does not tend to any limit even though the sample size is enormous and even though the series which it applies is presumably stationary’, (Mandelbrot, 1963, p. 395).

For both the FT-ALL and SP500 the recursive sample variance (shown in the figures below) seems to be time-varying. This evidence does not necessarily imply that it is non-convergent or infinite.



**Figure 3.** RLS of the FT-ALL variances.



**Figure 4.** RLS of the SP500 variance.

Mandelbrot (1963, 1964) proposed the Stable Paretian family of distributions (Levy, 1925) as being more appropriate in capturing the empirical regularities of symmetry, leptokurticity and infinite variance:

'independence in successive increments [could be] combined with a distribution other than the Gaussian' (Mandelbrot, 1964, p. 243).

The Stable Paretian family is usually specified using the characteristic function given by:

$$\log \phi(t) = i\delta t - \gamma |t|^\alpha \cdot \left[ 1 + i\beta \left( \frac{t}{|t|} \right) \tan \left( \frac{\pi\alpha}{2} \right) \right].$$

This family is quite flexible, given that it is described by four parameters;  $\alpha$  is the Pareto's exponent which determines the peakedness, ( $0 < \alpha \leq 2$ ), and smaller  $\alpha$  gives thicker tails;  $\beta$  measures the skewness, ( $-1 \leq \beta \leq 1$ ) and for  $\beta = 0$  the distribution is symmetric;  $\gamma$  is the scale parameter ( $-1 < \gamma < 1$ ); and finally  $\delta$  is the location parameter. Given this flexibility, the empirical regularities of leptokurticity, symmetry and infinite variance can be adequately modelled using this family. In addition, this family is closed (or stable) under linear transformations (e.g. summations) of IID random variables. The stability or invariance property implies that each stable distribution has an 'index of stability' (shape parameter) which remains the same regardless of the scale (sampling interval) adopted. Originally the stability concept was associated with the operations over IID random variables. However, due to this property, IID assumptions can be relaxed as a result of the CLT for non-ID random variables in the domain of attraction of the stable laws and the CLT for dependent random variables (e.g. Mittnik and Rachev, 1993a, b).

The use of the Pareto-Levy family of distributions in modelling returns was later called into question on several grounds. A number of studies find that the invariance property of the characteristic exponent is violated (e.g. Officer, 1972; Blattberg and Gonedes, 1974; Akgiray and Booth, 1988). These authors show that the characteristic exponent does not remain constant when going from weekly to monthly returns. Moreover, Praetz (1972) points out that the distribution functions of this family are unknown, except in several special cases, and that the estimation methods for their parameters are not very satisfactory. Furthermore, DuMouchel (1973) notes that the rate of convergence to the stable law may be very slow, implying that the stability property may not be very useful in practice. Lastly, Mandelbrot (1963) and Fama (1965) note that the infinite variance characteristic of Stable Paretian distributions makes statistical techniques based on the traditional asymptotic theory of finite-variance distributions, inappropriate. Moreover, financial analysis relies on the existence of both the mean and variance representing return and risk. The fact that some members of this family do not have a variance renders them inconsistent with most theoretical financial models.

Implicit in Mandelbrot's observations relating to the stylized facts, and explicit in his Stable Paretian model, is the assumption of temporal independence of returns. The author often notes that '*...successive daily changes of log prices are independent*' (Mandelbrot, 1964, p. 243), and in specifying his model he substitutes the Gaussian distribution (in Bachelier's model) with the Stable Paretian distribution. Hence, the probability model of the stock returns process is still a marginal rather than a conditional model. At the end of his paper, however, Mandelbrot (1963) observes that his model does not capture the observed alternation of small and big changes in fluctuations:

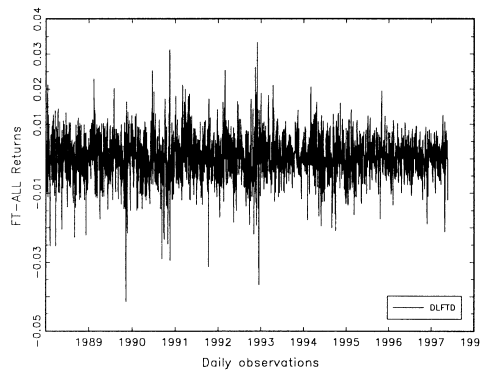
'Large price changes are not isolated between periods of slow change; they rather tend to be the result of several fluctuations, some of which 'overshoot' the final change. Similarly, the movement of prices in periods of tranquillity seem to be smoother than predicted by my process. In other words, large changes tend to be followed by large changes — of either sign — and small changes tend to be followed by small changes', (ibid. p. 418).

Along the same lines, Fama (1970) shows that:

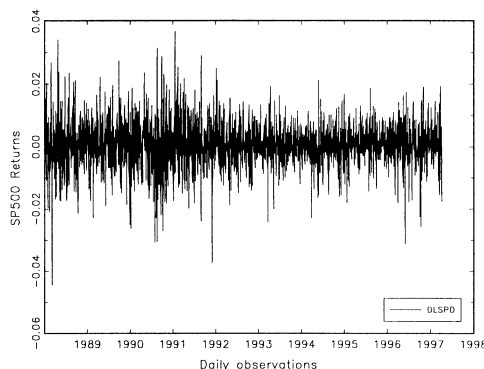
'...large price changes are followed by large price changes, but of an unpredictable sign. This suggests that important information can not be evaluated immediately', (ibid. p. 396).

The time-plots of the returns for the market indices considered (in Figures 5 and 6) provide information about the temporal behaviour of the processes. It is shown that large and small changes (of either sign) in returns tend to be clustered together over time.

This was later diagnosed as second-order temporal dependence (see, for example, Engle, 1982; Brock *et al.*, 1991), or volatility clustering of the returns process. This evidence implies that both the Random Walk and Heterogeneous 'Random Walk' models specified in the 1950s were inappropriate. Nevertheless, this regularity, although observed, cannot be modelled by the Stable Paretian



**Figure 5.** t-plot of FT-ALL returns.



**Figure 6.** t-plot of SP500 returns.

models. The evidence of second-order dependence in returns has also important implications for the infinite variance inferred by Mandelbrot (1963). The time-dependent or divergent unconditional recursive variance does not necessarily imply that this moment is infinite. This behaviour might be due to higher-order temporal dependence and/or time-heterogeneity. The presence of such features is well documented in the literature on speculative prices (e.g. Bollerslev *et al.*, 1994).

The literature in the 1970s attempted to deal with the presence of leptokurticity and the divergence of the second moment, largely ignoring the non-linear temporal dependence. DuMouchel (1973) proposes a mixture of independent Normal distributions, whereas Praetz (1972) and Blattberg and Gonedes (1974) consider two continuous mixture-of-Normal distributions models. They assume that the variance of the Normal distribution follows an inverted gamma distribution, the resulting posterior distribution is Student's  $t$ . They show that the non-Normal Stable Paretian distribution is obtained when the variance of the Normal distribution follows a strictly positive stable distribution with characteristic exponent between 0 and 1. Hence, Blattberg and Gonedes (1974) indicate that the Student's  $t$  distribution provides a better empirical description of stock returns than the Stable Paretian model (also see Hagerman, 1978; Lau *et al.*, 1990; Tucker, 1992). Some other recent studies (e.g. Akgiray and Booth, 1988; Phillips and Loretan, 1994) have found evidence against the Stable model. The Paretian tail index is estimated from the tail observations, using the Pareto distribution and its generalization. The results of these studies show a tail index significantly greater than 2. McCulloch (1997) criticises this inference advocating that the previous studies overlook the implications of the intermediate value theorem. Given the moderate sample sizes often used, tail-index estimates in excess of 2 are to be expected for IID Stable samples with exponent as low as 1.65. Therefore, an exponent greater than 2 is consistent with a stable distribution for asset returns. Maximum Likelihood estimation for the full-sample is recommended for tail-index estimators if the distribution is assumed to be Stable. Nevertheless, this analysis largely relies on the IID assumption of returns.

Recently, Mittnik and Rachev (1993a,b) propose an alternative distribution in the Pareto family, the double Weibull distribution, generated from geometric sums of IID random variables. Their empirical results show that the Weibull distribution dominates the Stable distribution in terms of fit (for the S&P500 daily and monthly returns, 1982–87). Their work has been criticized for the following reasons: (i) Although numerous distributions have been found to have an adequate *fit* for speculative price data, the essence of a statistical model is to capture the data generating mechanism and to give rise to reliable statistical inference results (Phillips, 1993). (ii) There is no empirical support of the covariance stationarity assumption for stock returns (e.g. Phillips and Loretan, 1994). (iii) The IID assumption ignores the conditional density function that incorporates important systematic, temporal dependence information of the returns process (Diebold, 1993; Baillie, 1993).

The Mandelbrot era ended in the early 1980s with the Normality assumption of the Random Walk model replaced by a variety of leptokurtic distributions. Although the presence of second-order dependence was inconsistent with the Heterogeneous Random Walk model, this issue was left unaddressed.

### 3.3. *The Efficient Market Hypothesis*

In the 1970s the literature faced the challenge of respecifying the theoretical model consistent with the notion of market efficiency. By the mid 1960s Mandelbrot (1967), Samuelson (1965) and Fama (1970) acknowledged the fact that a market could be efficient even if returns were not an IID process. What was really required for efficiency was the lack of any rule of buying and selling that has an expected return greater than the average return of the market. In other words, the game of speculation should be fair or (excess) returns should follow a martingale difference process. The statistical definition of a fair game implies that the conditional expectation of returns at time  $t$ , based on all past information on returns, should be zero:

$$E(r_t | \sigma(r_{t-1}, \dots, r_1)) = 0, \quad t \in \mathbb{T}. \quad (8)$$

Equivalently, in an efficient market the best forecast for today's prices is yesterday's price:

$$E(p_t | \sigma(p_{t-1}, \dots, p_1)) = p_{t-1}, \quad t \in \mathbb{T}. \quad (9)$$

From the economic theory viewpoint, the switch from the random walk to the less restrictive martingale model was based on the need to link the behaviour of prices with more fundamental economic quantities, the role of anticipation and expected utility and the concept of an efficient market. The random walk model implies that stock prices were exempt from the supply and demand laws and look like the casino or musical chairs that Keynes (1936) chose as a metaphor for the stock market. Hence, the random walk model was inconsistent with the traditional theory of competition. Moreover, the probabilistic independence



between successive price increments, as assumed by the random walk model, is quite restrictive when looked upon from a broad class of optimizing problems. Mandelbrot (1966) demonstrates that the martingale model may arise in the context of a ‘fundamental analysis’ in which the price attempts to follow ‘something that can be described as value’.

Samuelson (1965) also provided a formal proof of a theorem which states that properly anticipated prices fluctuate randomly iff some specific restrictions hold. These restrictions require agents to have common and constant time preferences, have common probabilities and be risk neutral. Samuelson restricts his analysis to the case of ‘futures pricing’, but his arguments can be extended to the spot-price case as well.

The martingale process has become a powerful tool in probability and statistics (see Doob, 1953), and modern theories of asset prices. For example, once asset prices are properly adjusted for risk, the martingale property does hold (see Lucas, 1978; Harrison and Kreps, 1979) and the combination of this risk-adjustment and the martingale property has led to the pricing of complex financial instruments such as options, swaps and other ‘derivatives’ (Merton, 1990). The Martingale formulation (9) has also certain advantages from the statistical viewpoint. Note that the literature of the Bachelier-Kendall era viewed model (3) as a system driven by the IID process for  $\{r_t, t \in \mathbb{T}\}$ . The probabilistic structure of the process  $\{p_t, t \in \mathbb{T}\}$  was indirectly determined by contemplating (3) from right to left. The Martingale formulation (9) reverses the role and considers  $\{p_t, t \in \mathbb{T}\}$  as the primary process, contemplating (3) from left to right as an orthogonal decomposition of the form:

$$p_t = E(p_t | \sigma(p_{t-1}, \dots, p_1)) + r_t, \quad t \in \mathbb{T}.$$

The process  $\{r_t, t \in \mathbb{T}\}$  constitutes the unmodelled component defined by:

$$r_t = p_t - E(p_t | \sigma(p_{t-1}, \dots, p_1)), \quad t \in \mathbb{T},$$

which is by ‘design’ a Martingale difference process, denoted by  $r_t \sim \text{MD}(0)$ , with properties:

$$\left. \begin{aligned} \text{(i)} \quad & E(|r_t|) < \infty, \\ \text{(ii)} \quad & E(r_t | \sigma(r_{t-1}, \dots, r_1)) = 0, \end{aligned} \right\} t \in \mathbb{T}. \tag{10}$$

It is well known that the probabilistic structure of a Martingale difference process is much weaker than either an IID (with moments up to order two) or an Independent but non-Identically distributed process with  $t$ -dependent variance. The IID assumption may alternatively be expressed as:

$$D(r_t | r_{t-1}, \dots, r_1; \theta) = D(r_t; \theta), \quad \text{for all } t \in \mathbb{T}. \tag{11}$$

In terms of the conditional moments (assuming they exist), (11) can be expressed in the form of:

$$E(r_t^k | \sigma(r_{t-1}, \dots, r_1)) = E(r_t^k), \quad k = 1, 2, \dots, \quad \text{for all } t \in \mathbb{T}. \tag{12}$$

In contrast, a Martingale difference process assumes the existence of the first moment only (i.e. for  $k = 1$  in (12)). This point was also made by Mandelbrot as far back as 1966, when he first introduced the martingale model for stock prices. He explicitly states that:

‘It should be stressed that the distribution of  $r_{t+T}$ , conditioned by known values of  $r_t$  and of  $r_{t-i}$ , may very well depend upon the past values  $r_{t-i}$ : the expectation alone is unaffected by  $r_{t-i}$ ’ (ibid, p. 244).

### 3.4. Higher order martingale difference

One can extend a martingale difference process to include higher conditional moments (see Spanos, 1999). In particular, higher order dependence may arise through higher conditional moments (assuming they exist):

$$\left. \begin{aligned} E(r_t | r_{t-1}, \dots, r_1) &= 0, \\ E(r_t^k | r_{t-1}, \dots, r_1) &= g_k(r_{t-1}, r_{t-2}, \dots, r_1), \quad \text{for } k = 2, 3, \dots, \end{aligned} \right\} t \in \mathbb{T}. \quad (13)$$

A second order martingale difference process is defined by (10) and the conditional variance:

$$E(r_t^2 | \sigma(r_{t-1}, \dots, r_1)) = \xi(r_{t-1}, \dots, r_1). \quad (14)$$

It is interesting to note that Mandelbrot (1964, 1969) has alluded to such a possibility:

‘... no policy exists for buying and selling that has an expected return greater than the average return of the market. On the other hand, the martingale model does allow the actual distribution of  $[p_t - p_{t-1}]$  to depend on past and present prices, and therefore it does not deny that past and present prices can serve in the selection of portfolios of different desired degrees of riskiness’, as represented by the second conditional moment (Mandelbrot, 1969, p. 227).

Note that in the case of a homoskedastic conditional variance:

$$E(r_t^2 | r_{t-1}, \dots, r_1) = \sigma^2, \quad t > 2,$$

a second order martingale difference process becomes an *innovation* process.

The most direct empirical tests on the appropriateness of (13) amount to determining whether there exist significant conditional mean effects (LeRoy, 1989). As noted above, however, the martingale difference model does not preclude any significant effects in higher order conditional moments (e.g. the conditional variance). Indeed, the temporal dependence puzzle, observed by Mandelbrot (1963) and Fama (1970), is consistent with the martingale difference model since the latter can accommodate any non-linear dynamic effects via higher-order conditional moments. A higher-order martingale difference model can also be considered as a useful framework for specifying a model with

conditional skewness and kurtosis (e.g.  $k = 3, 4$ ), where there are a function of the conditioning information set (see for instance, Hansen, 1994).

**4. The dynamic volatility era (1980– present)**

By the early 1980s it was clear that the existing empirical models could not account for the non-linear temporal dependence exhibited by speculative price data. To illustrate that, let us consider the temporal dependence structure of the SP500 and FT-ALL stock returns series. Table 2 presents the results from three alternative temporal dependence tests on the residuals of an AR(1) model for each series. The Ljung-Box (1979) (LB) tests the null hypothesis that the residuals is an uncorrelated process, whereas the McLeod-Li (1981) (ML) and the BDS (Brock *et al.*, 1987, 1991, 1996) test for the existence of higher order dependence.

The results show that the SP500 and FT-ALL returns do not reject the martingale difference hypothesis. Both the McLeod-Li and BDS tests present evidence of non-linear dependence in both series. Hence, there is no empirical support of the random walk model with independent increments.

Second-order dependence may be modelled via a quadratic dynamic heteroskedastic function. The first attempt in the direction of modelling the conditional variance was made by Engle (1982), who proposed the AutoRegressive Conditionally Heteroskedastic (ARCH) model (note 1) specified in terms of the first two conditional moments:

$$E(r_t | \mathbf{R}_{t-1}^p) = g_t(r_{t-1}, r_{t-2}, \dots, r_{t-p}), \tag{15}$$

$$Var(r_t | \mathbf{R}_{t-1}^p) = E\{[r_t - E(r_t | \mathbf{R}_{t-1}^p)]^2 | \mathbf{R}_{t-1}^p\} = h_t(r_{t-1}, \dots, r_{t-p}). \tag{16}$$

where  $\mathbf{R}_{t-1}^p := \{r_{t-1}, r_{t-2}, \dots, r_{t-p}\}$ .

**Table 2.** Temporal Dependence tests for US & UK Stock Market Returns

	LB(5)	LB(10)	LB(15)	ML(5)	ML(10)	ML(15)
SP500	4.960 (0.421)	14.129 (0.167)	24.696 (0.054)	35.553 (0.000)**	76.692 (0.000)**	116.096 (0.000)**
FT-ALL	8.224 (0.144)	11.805 (0.298)	22.464 (0.096)	96.202 (0.000)**	131.478 (0.000)**	205.877 (0.000)**
	BDS(1, 2)	BDS(0.25, 5)	BDS(0.25, 10)	BDS(0.5, 2)	BDS(0.5, 5)	BDS(0.5, 10)
SP500	89.644**	6.264*	8.980**	-0.091	0.088	1.593
FT-ALL	59.798**	7.020**	8.984**	9.409**	9.884**	7.719**

Notes:

1. LM = Lagrange Multiplier test, LB = Ljung-Box test, ML = McLeod-Li test, BDS = Brock *et al.* test. The numbers in brackets refer to p-values.
2. (\*\*) refer to rejection of the null hypothesis at the 5% significance level.

This model is equivalent to an extended martingale difference process if  $E(r_t | \mathbf{R}_{t-1}^p) = 0$ . Given this, second-order stationarity holds if:

$$\text{Var}(r_t) = E\{\text{Var}(r_t | \mathbf{R}_{t-1}^p)\} = \text{constant}.$$

A number of models with dynamic heteroskedasticity have been proposed in the literature using alternative parameterizations of  $\text{Var}(r_t | \mathbf{R}_{t-1}^p)$ . Initially, Engle (1982) expressed the conditional variance in the ARCH model as a  $p^{\text{th}}$  order weighted average of past squared disturbances:

$$\text{var}(r_t | \mathbf{R}_{t-1}^p) = a_0 + a_1 e_{t-1}^2 + \cdots + a_p e_{t-p}^2, \quad (17)$$

$$\text{where } e_t = r_t - E(r_t | \mathbf{R}_{t-1}^p), \quad a_0 > 0, \quad a_i \geq 0, \quad i = 1, \dots, p.$$

Early empirical studies using the ARCH model found that the number of lags in (17) was rather sizeable. Therefore, in their attempt to reduce the number of parameters in ARCH, they imposed ad hoc linearly declining lag structures such as:

$$a_i = \frac{\alpha(p+1-i)}{p(p+1)},$$

for some  $\alpha > 0$ , (Engle, 1982). Motivated by the problem of long lags an alternative parameterization of the dynamic conditional variance was proposed by Bollerslev (1986) known as the Generalized ARCH. Using the analogy with ARMA( $p, q$ ) models, Bollerslev specified the conditional variance to be:

$$\text{Var}(r_t | \mathbf{R}_{t-1}^p) = a_0 + \sum_{j=1}^p b_j h_{t-j}^2 + \sum_{i=1}^q \alpha_i e_{t-i}^2, \\ \alpha_0 > 0, \quad 0 \leq \alpha_i < 1, \quad i = 1, \dots, q, \quad 0 \leq b_j < 1, \quad j = 1, \dots, p, \quad (18)$$

known as the GARCH( $p, q$ ) model. These models provide a more parsimonious representation than a higher-order ARCH model, and thus they are easier to identify and estimate.

In the ARCH model, for the second moment to exist, we need to impose the additional restriction:  $(1 - \sum_{i=1}^p a_i) > 0$ , (Engle, 1982). For the existence of higher moments, such as the kurtosis coefficient, more stringent restrictions are often needed. The GARCH model imposes similar restrictions for the variance to exist, given by  $(\sum_{i=1}^q \alpha_i + \sum_{j=1}^p b_j) < 1$ . Of particular interest in the empirical literature is the case for which the sum of the variance coefficients is one, i.e.  $(\sum_{i=1}^q \alpha_i + \sum_{j=1}^p b_j) = 1$ . This gives rise to what is known as the Integrated GARCH (IGARCH) model (Engle and Bollerslev, 1986a, b). Volatility in such processes is highly persistent. Although the IGARCH(1, 1) process is often seen as analogous to a linear random walk type conditional variance, Nelson (1990) shows that this analogy is misleading. A linear random walk is nonstationary in two respects. First, it has no stationary distribution implying that the process is

strictly nonstationary. Second, the first two unconditional moments do not exist i.e. it is second-order non-stationary. On the other hand, the IGARCH(1,1) model is paradoxical in the sense that the implied process is strictly stationary even though its distribution does not have the necessary second (unconditional) moments. Note that such integrated models agree with Mandelbrot's assertion about an infinite second moment.

In Tables III and IV we present the estimation and misspecification test results of the ARCH( $p$ ) and GARCH( $p, q$ ) models, respectively. The misspecification test results suggest that the Normal ARCH models do not account for the leptokurticity of the conditional distribution (as shown by the D'Agostino kurtosis test) and do not capture all the second-order dependence in higher-order lag lengths (as shown by the ARCH and McLeod-Li tests). Therefore, GARCH(1,1) models are estimated which present a more parsimonious representation for stock returns than the ARCH( $p$ ) models. Generally, the Normal GARCH(1,1) models seem to perform better on statistical adequacy grounds than the ARCH( $p$ ) models, since the former capture all the non-linear

**Table 3.** Estimation results of GARCH-type models

	ARCH		N-GARCH		t-GARCH	
	SP500	FT-ALL	SP500	FT-ALL	SP500	FT-ALL
$\gamma_0$	0.049 (0.020)	0.039 (0.020)	0.056 (0.019)	0.048 (0.019)	0.075 (0.016)	0.058 (0.018)
$\gamma_1$	0.045 (0.023)	0.101 (0.023)	0.027 (0.021)	0.096 (0.022)	0.009 (0.019)	0.074 (0.021)
$a_0$	0.921 (0.033)	0.885 (0.031)	0.003 (0.001)	0.038 (0.011)	0.005 (0.003)	0.020 (0.014)
$a_1$	0.081 (0.026)	0.103 (0.023)	0.015 (0.003)	0.075 (0.013)	0.026 (0.006)	0.047 (0.017)
$b_1$	—	—	0.981 (0.004)	0.888 (0.021)	0.971 (0.007)	0.932 (0.029)
$\nu$					4.708 (0.496)	8.263 (1.189)
$\ln L$	-3316.9	-3296.4	-3222.3	-3229.3	-1761.7	-1839.8
$a_1 + b_1$	—	—	0.996	0.963	0.997	1.010
$H_0:$						
$a_1 + b_1 = 1$	—	—	3.326 [0.068]	4.077 [0.043]	0.959 [0.327]	2.131 [0.144]

Note: The  $\gamma_i$ 's refer to the linear AR model parameters. The rest of the coefficients refer to the GARCH model parameters. The numbers in brackets refer to standard errors. The numbers in square brackets refer to  $p$ -values

**Table 4.** Misspecification Tests Results of GARCH-type Models

	ARCH		N-GARCH		t-GARCH	
	SP500	FT-ALL	SP500	FT-ALL	SP500	FT-ALL
D'Ag. Sk.	-9.381 (0.000)**	-1.320 (0.093)	-10.519 (0.000)**	-2.004 (0.023)	-10.010 (0.000)**	-4.130 (0.000)**
D'Ag. Kurt.	16.736 (0.000)**	13.105 (0.000)**	16.008 (0.000)**	8.111 (0.000)**	16.429 (0.000)**	6.977 (0.000)**
KG2	2.235 (0.135)	0.909 (0.340)	3.898 (0.048)	0.079 (0.778)	3.849 (0.050)	0.002 (0.969)
ARCH(1)	0.176 (0.675)	0.030 (0.862)	0.002 (0.964)	0.008 (0.931)	0.088 (0.767)	0.524 (0.469)
ARCH(5)	8.629 (0.000)**	5.922 (0.000)**	0.549 (0.739)	9.642 (0.000)**	0.435 (0.825)	0.871 (0.499)
ARCH(10)	9.746 (0.000)**	5.597 (0.000)**	1.321 (0.213)	6.314 (0.000)**	1.198 (0.287)	0.509 (0.885)
ARCH(15)	6.352 (0.000)**	5.492 (0.000)**	1.214 (0.253)	6.954 (0.000)**	1.145 (0.309)	1.126 (0.326)
LM (1)	1.152 (0.283)	0.021 (0.885)	0.144 (0.704)	0.896 (0.344)	0.511 (0.475)	0.018 (0.892)
LM (5)	1.124 (0.345)	1.602 (0.156)	1.055 (0.383)	2.241 (0.048)	1.190 (0.311)	1.382 (0.228)
LM (10)	1.489 (0.137)	1.747 (0.156)	1.095 (0.362)	1.596 (0.102)	1.143 (0.325)	1.099 (0.358)
LB (10)	11.300 (0.335)	6.912 (0.227)	5.184 (0.394)	13.696 (0.187)	12.549 (0.249)	10.427 (0.404)
LB (15)	20.034 (0.170)	15.803 (0.105)	9.697 (0.467)	23.825 (0.068)	24.211 (0.062)	17.426 (0.294)
LB (20)	22.808 (0.298)	23.348 (0.077)	26.020 (0.165)	33.424 (0.030)	29.433 (0.079)	24.777 (0.210)
ML (10)	55.426 (0.000)**	31.965 (0.000)**	4.672 (0.912)	82.957 (0.000)**	3.974 (0.948)	5.077 (0.886)
ML (15)	82.474 (0.000)**	68.798 (0.000)**	8.383 (0.907)	148.724 (0.000)**	7.301 (0.949)	16.847 (0.328)
ML (20)	90.789 (0.000)**	109.926 (0.000)**	9.252 (0.979)	180.398 (0.000)**	8.208 (0.990)	21.041 (0.394)

Notes:

1. D'Ag. Sk. and Kurt. = D'Agostino Skewness, Kurtosis test. KG2 = Linearity test (Kolmogorov-Gabor polynomials). ARCH test. LM = Lagrange Multiplier test. LB = Ljung-Box test, ML = McLeod-Li test.
2. The numbers in brackets refer to  $p$ -values.
3. (\*\*\*) refer to rejection of the null hypothesis at the 5% significance level.

dependence left unmodelled by the latter. Nevertheless, the empirical results suggest that the Normal GARCH models do not seem to account for conditional leptokurticity (as shown by the rejection of the D'Agostino kurtosis test). Hence, although the unconditional distribution of the errors in the Normal GARCH( $p, q$ ), has fatter tails than the Normal distribution it does not adequately account for the leptokurtosis in the SP500 and FT-ALL returns data.

To summarize, the stochastic process for returns which satisfies certain theoretical requirements and captures the stylized facts exhibited by data on stock returns, can be adequately described by an extended martingale difference process with dynamic heteroskedasticity effects. Clearly, such a process cannot be Gaussian. In view of this, alternative conditional distributions,  $D(r_t | \mathbf{R}_{t-1}^p)$ , have been proposed. Initially, Engle (1982) assumed conditional Normality, with a leptokurtic unconditional distribution. In the empirical studies that followed, it was realized that the resulting unconditional distribution did not account for all the leptokurtosis in the observed data (see inter alia Bollerslev, 1987). Bollerslev replaced the Normal distribution of the error with that of a  $t$ -distribution of the form:

$$D(e_t; \theta) = \frac{\Gamma[\frac{1}{2}(\nu + 1)]}{\Gamma[\frac{1}{2}\nu]} [(\nu - 2)h_t^2]^{-\frac{1}{2}} \left[ 1 + \frac{e_t^2}{(\nu - 2)h_t^2} \right]^{-\frac{1}{2}(\nu + 1)}. \tag{19}$$

As shown in McGuirk *et al.*, (1993) and Spanos (1994), this formulation does not represent the conditional Student's  $t$   $D(r_t | \mathbf{R}_{t-1}^p)$ , derived from the joint distribution  $D(r_1, r_2, \dots, r_T; \psi)$ . We proceed to estimate and test the  $t$ -GARCH(1, 1) model (Bollerslev, 1987) for the SP500 and FT-ALL stock returns data. The results in Table III show that this model can account for leptokurtosis more adequately than the Normal GARCH model. Similarly, the conditional variance coefficients are significant and the misspecification tests show that it adequately captures second-order dependence of the SP500 series.

Another distributional assumption, proposed by Nelson (1991), results to the Exponential GARCH (EGARCH) model which accounts for the empirical evidence that stock returns are negatively correlated with changes in return volatility (Black, 1976). This effect suggests that volatility tends to rise in response to 'bad news' (when excess returns are lower than expected) and fall in response to 'good news' (when excess returns are higher than expected). The reason for this behaviour of returns is financial and operating leverage (Nelson, 1991). The GARCH models disregard this regularity since they assume that only the size and not the sign of unanticipated excess returns determines the conditional variance. On the other hand, the EGARCH model responds asymmetrically to positive and negative effects on the conditional variance:

$$\ln(h_t) = \alpha_t + \sum_{k=1}^q \beta_k (\theta r_t + \gamma [\eta_{t-k} - E | \eta_{t-k} | ]) + \sum_{k=1}^p \delta_k \ln(h_{t-i}), \quad \beta_1 = 1, \tag{20}$$

where  $\eta_t = r_t/\sigma_t$ . The term  $\gamma[\eta_{t-k} - E|\eta_{t-k}|]$  represents a magnitude effect (which is the relevant effect in GARCH models) and the term  $\theta r_t$  represents the sign effect which can account for the regularity observed by Black (1976). Note that unlike the GARCH( $p, q$ ) model the EGARCH model does not impose any nonnegativity constraint on the parameters  $\beta_k$  and  $d_k$  of the conditional variance, since the latter model deals with the *logarithm* of the variance. (Also see note 2.)

The non-Normality in the distribution of returns is introduced by assuming that  $r_t$  is drawn from the Generalized Error Distribution density (or equivalently the exponential power distribution):

$$D(e_t; \theta) = \frac{v \exp[-(\frac{1}{2}) |e_t/\lambda|^v]}{\lambda 2^{(1+1/v)} \Gamma(1/v)}, \quad e_t \in \mathbb{R}, \quad 0 < v \leq \infty,$$

where  $\Gamma(\cdot)$  is the gamma function and  $\lambda := [2^{(-2/v)} \Gamma(1/v) / \Gamma(3/v)]^{1/2}$ .

Stochastic volatility models is another class of parameterizations that have been utilized to capture the mechanism of speculative prices (for a recent survey, see Ghysels *et al.*, 1996). Their continuous time aspect implies that they are closely related to theoretical models of finance as compared to GARCH-type models. In addition, stochastic volatility models regard volatility as being driven by a process separate from returns per se. They are based on an EGARCH functional form but the process driving the conditional variance is a NID(0, 1) process, that is independent of  $r_t$ . Such a model has the format:

$$\log \sigma_t^2 = \alpha_0 + \beta_1 \log \sigma_{t-1}^2 + \sigma_n \eta_t.$$

A variant of the stochastic volatility model that has been extensively used by Hamilton (1990) is to relate  $\sigma_t^2$  to an unobserved state variable  $z_t$  that can take either the value 0 or 1, with this variable evolving according to a first order Markov process. This model differs from the standard stochastic volatility model in two ways: The innovations driving the conditional variance are a discrete random variable with four states and the conditional variance of the innovations depends on  $\sigma_{t-1}^2$ .

An attempt to link the GARCH and stochastic volatility models has arisen as a result of the high degree of estimated persistence observed after fitting GARCH or stochastic volatility models. Parameter instability has been regarded as one of the reasons for volatility persistence (e.g. Lamoureux and Lastrapes, 1990) and Markov switching ARCH model (Hamilton and Susmel, 1994; Cai, 1994) have been specified as a response.

To sum up, since the early 1980s the literature dealt with the stylized facts of non-linear dynamics and leptokurticity of speculative price data by postulating functional forms for the conditional variance and associating that with *conditional* leptokurtic distributions. The particular functional forms have been justified in terms of apparent analogies with the corresponding ARMA type models for the



conditional mean. Although plausible, these functional forms are ad hoc and their only justification lies with their empirical validity. Statistical adequacy is a necessary justification for a statistical model; Yet the theoretical justification for conditional variance functional forms is still an open question.

One of the main messages drawn from the discussion so far is that the empirical literature has concentrated at certain points in time on postulating statistical models to deal with some empirical regularities, giving more emphasis in one or another such regularity. During the Bachelier era the emphasis was placed on modelling first-order temporal dependence. The Mandelbrot era emphasized the distributional aspect of modelling having interpreted the erratic behaviour of the sample recursive variance as a symptom of an infinite unconditional variance. The dynamic volatility era emphasized the modelling of the conditional heteroskedasticity as a way to capture the higher order temporal dependence. The other aspects of empirical modelling, such as the distribution, are seen only through the prism of dynamic heteroskedasticity. Nevertheless, all three probabilistic aspects, namely distribution, dependence and heterogeneity, are equally important for specifying the relevant statistical model.

An all encompassing approach which attempts to deal with all the observed empirical regularities has been made in the context of the Probabilistic Reduction approach; see Spanos (1986,1995). This approach views models as reductions from the joint distribution of the observable random variables involved. In the case of a single stochastic process  $\{r_t, t \in \mathbb{T}\}$  the postulated statistical models can be viewed as simplifications of the joint distribution via the imposition of probabilistic assumptions which restrict the sequential conditioning:

$$D(r_1, r_2, \dots, r_T; \psi) = D(r_1; \varphi_1) \prod_{t=2}^T D(r_t | \mathbf{R}_{t-1}^0), \quad \forall (r_1, r_2, \dots, r_T) \in \mathbb{R}^T.$$

The idea is to choose the reduction assumptions from the three basic categories in (1) in such a way so as to allow for the possibility of accommodating all the observed empirical regularities.

An example of a statistical model specified in the context of this approach is the Student's  $t$  AutoRegressive model with dynamic heteroskedasticity (Spanos, 1991,1994). By assuming that the joint distribution is Student's  $t$  and stationary and allowing for weak dependence, this model accounts (i) for joint and marginal leptokurticity, (ii) second order temporal dependence, (iii) the erratic behaviour of the sample recursive variance. The reduction process ensures the internal consistency of the model assumptions and requires no additional parameter restrictions, such as those needed for the GARCH family of models. This model takes the form:

$$r_t = \gamma_0 + \sum_{i=1}^{\ell} \gamma_i r_{t-i} + u_t, \quad \ell > 0, \quad t \in \mathbb{N}, \tag{21}$$

$$\omega_t^2 = \left( \frac{\nu}{\nu + t - 3} \right) \sigma^2 \left( 1 + \sum_{j=1}^{t-1} \sum_{i=1}^{t-1} \{\delta_{ij}[r_{t-i} - \mu][r_{t-j} - \mu]\} \right), \quad (22)$$

where,  $\delta_{ij} = 0$  for all  $|i - j| > p$ ,  $u_t := r_t - E(r_t | \sigma(\mathbf{R}_{t-1}^0)) \sim St(0, \omega_t^2; \nu)$ ,  $\omega_t^2$  is the conditional variance,  $\mu = E(r_t)$ ,  $\nu > 2$  is the degrees of freedom, and  $\sigma(\mathbf{R}_{t-1}^0)$  denotes the sigma-field (the conditioning information set) generated by  $\mathbf{R}_{t-1}^0 := (r_{t-1}, r_{t-2}, \dots, r_1)$ . From (21) the conditional mean is a linear function of the conditioning variables and resembles the simple Normal AR( $\ell$ ) process. Nevertheless, the conditional variance (22) is heteroskedastic and thus able to encapsulate the observed second-order dependence. Specifically the conditional variance is a quadratic function of all the past conditioning information, but is parameterized with only  $p + 1$  unknown  $\delta_j$ 's.

There are numerous recent developments for modelling speculative prices. We choose to focus on only two of those that are related with the dynamic heteroskedastic volatility models: The evidence of fractional integration and long memory in returns and the modelling of higher order conditional moments. Recent research examines whether speculative prices are fractionally integrated and can be described as long memory processes. A fractionally differenced process  $\{p_t, t \in \mathbb{T}\}$ , satisfies the following difference equation:

$$(1 - L)^d p_t = \varepsilon_t, \quad \varepsilon_t \sim IID(0, \sigma_\varepsilon^2)$$

and  $d$  may be noninteger powers that define a fractionally series of order,  $d$ ,  $I(d)$ . For the early and significant work on ARFIMA processes see Granger (1981), Granger and Joyeux (1980), and Hosking (1981). When  $d \in (-0.5, 0.5)$ ,  $p_t$  is stationary and invertible and exhibits a unique kind of dependence that depends on the sign of  $d$ : For  $0 < d < 0.5$ , the autocorrelation coefficients are all positive and decay at a hyperbolic rate, so that the process belongs to the class of long-memory processes (see Hurst, 1951, for the origin of interest in long memory processes in the hydrology field) (see note 5). For  $-0.5 < d < 0$ , the sum of absolute values of the process autocorrelations tends to a constant; hence it has a short memory. For a recent comprehensive survey of long memory and fractionally integrated processes that reviews the models developed under this class, various tests, estimation methods for fractional integration and a wide range of applications, see Baille (1996).

The original statistical measurement of long memory due to Hurst (1951) and used by Mandelbrot (1972) is the rescaled ( $R/S$ ) statistic, defined as:

$$R_T = \max_{0 \leq j \leq T} \left\{ \sum_{j=1}^T (y_j - \bar{y}) \right\} - \min_{0 \leq j \leq T} \left\{ \sum_{j=1}^T (y_j - \bar{y}) \right\}$$

where  $R$  = range,  $\bar{y}$  is the sample mean. Hurst (1951) showed that:

$$p \lim \{(T)^{-H} (R_T/S_T) = \text{constant}, \quad (23)$$

where  $S_T$  is the sample standard deviation:

$$S_T = \left\{ \left( \frac{1}{T} \right) \sum_{j=1}^T (y_j - \bar{y}) \right\}^{1/2} .$$

Following (23) we obtain:

$$\log[E(R_T/S_T)] \approx \text{constant} + H[\log(T)],$$

and the Hurst coefficient,  $H$ , can be estimated either as the ratio of  $\log(R_T/S_T)/\log(T)$  or the regression coefficient of  $\log(R_T/S_T)$  on  $\log(t)$ , for different  $t$  (Baille, 1996, presents alternative methods of estimating  $H$ ). Since a short memory process would have  $H = 0.5$ , an estimated value of  $H > 0.5$  is interpreted as evidence of long memory. Lo (1991) modified the  $R/S$  statistic in the presence of short memory and heteroskedasticity. Applying the two statistics to CRSP stock returns indices, Lo finds that the original  $R/S$  statistic gives significant results whereas the modified  $R/S$  gives insignificant results. This constitutes evidence against the existence of long memory in the returns process.

Recently Ding *et al.* (1993) have shown the presence of long memory in higher moments of returns series. This has led to the development of Fractionally Integrated GARCH and EGARCH models (see Baille *et al.*, 1996 and Bollerslev and Mikkelsen, 1996, respectively) to capture long memory in volatility. It is found that the FIEGARCH model which allows for nonsymmetry of shocks provided a good representation of the SP500 volatility. Ding and Granger (1996) extend the concept of long memory processes and define generalized fractionally integrated models which may arise from aggregation, time-changing coefficient models and nonlinear models. They show that certain useful transformations of a returns process may produce long memory, even though the actual returns series has no such property. These transformations refer to absolute returns and power transformations of the daily SP500 stock market returns. Research in this area shows that econometric models are extended to allow for evidence of fractional integration. They seem to focus in either of the first two conditional moments. Although it would be interesting to integrate the two, the appropriate tests to distinguish between  $I(0)$ ,  $I(1)$  and  $I(d)$  is an equally important open area.

The analysis so far has centred on modelling the conditional mean and variance of asset returns. Higher-order conditional moments of a bivariate process  $\{w_t \equiv (y_t, x_t), t \in \mathbb{T}\}$ , (other than the mean and variance) may be needed to describe the conditional distribution. For instance, modelling the conditional skewness and kurtosis is important for efficient estimation, prediction and complete specification of asset pricing models. In Hansen (1994) the conditional density depends upon a low-dimensional 'parameter vector' and is allowed to vary as a function of the conditional variables. This generates a class of models called the Autoregressive Conditional Density (ARCD) models. Their distribution is given by:

$$D(y/a(x_t, \theta)) = \frac{d}{dy} P(y_t \leq y/x_t)$$

where  $\theta$  is a finite-dimensional parameter vector and  $a_t = a(x_t, \theta)$  is a low-dimensional time-varying parameter which fully describes the influence of  $x_t$  upon the conditional distribution. For example, it is assumed that  $a_t = (\mu_t, \sigma_t^2, \eta_t)$  where  $\mu_t$  is the conditional mean,  $\sigma_t^2$  is the conditional variance and  $\eta_t$  denotes the additional shape parameters that are time-variant.

A flexible density function is chosen that can generate a range of shapes (e.g. fat tails, skewness). In addition, a closed form density function is required for using Quasi-MLE. Hansen (1994) generalizes the conditional Student's  $t$  density function to allow for skewness. Hence the conditional distribution of the normalised error,  $z$ , is:

$$g(z/\eta, \lambda) = \begin{cases} bc \left( 1 + \frac{1}{\eta - 2} \left( \frac{bz + a}{1 - \lambda} \right)^2 \right)^{-(\eta + 1)/2} & z < -a/b \\ bc \left( 1 + \frac{1}{\eta - 2} \left( \frac{bz + a}{1 + \lambda} \right)^2 \right)^{-(\eta + 1)/2} & z \geq -a/b \end{cases}$$

where  $2 < \nu < \infty$ , and  $-1 < \lambda < 1$ . The constants  $a, b, c$  are given by:

$$a = 4\lambda \left( \frac{\eta - 2}{\eta - 1} \right), \quad b^2 = 1 - 3\lambda^2 - a^2, \quad c = \frac{\Gamma\left(\frac{\eta + 1}{2}\right)}{\sqrt{\pi(\eta - 2)}\Gamma\left(\frac{\eta}{2}\right)}.$$

This 'skewed Student's  $t$ ' distribution specializes to the traditional Student's  $t$  (with unit variance) if  $\lambda = 0$ . If  $\lambda > 0$ , the mode of the density,  $(-a/b)$ , is to the left of zero and the variable is skewed to the right, and vice-versa when  $\lambda < 0$ . Following the ARCH approach,  $\eta_t$  is also modelled as a function of lagged residuals:

$$\eta_t = n(e_{t-1}, e_{t-2}, \dots, e_1).$$

Parameter constraints allow  $\eta_t$  to lie within a region  $[L, U]$ , where the relationship between  $\eta_t$  and  $\lambda_t$  is given by the logistic transformation:

$$\eta_t = L + \frac{(U - L)}{1 + \exp(-\lambda_t)}$$

and  $\lambda_t$  is allowed to vary over the entire real line. The applications of the 'skewed Student's  $t$ ' conditional density show that the shape parameters are statistically significant at the 5% level. The approach adopted in this study points to a direction of extending statistical parameterizations to higher-order conditional

moments beyond the conditional variance. Yet, the question still remains open in bridging the gap between statistical and theoretical models for asset returns.

## 5. Conclusions

The review of the empirical modelling of speculative price data since the early 20th century suggests that the choice of statistical models can be best described as a dialectic process between empirical regularities and appropriate probabilistic assumptions. In view of the apparent leptokurticity exhibited by the histograms of such data the Normality distribution was replaced by other leptokurtic distributions. The detection of higher order temporal dependence in such data led to the abandonment of correlation as the sole form of dependence. This eventually led to the adoption of a martingale difference process with higher order dynamics. The erratic behaviour of the sample recursive variance, which was initially interpreted as evidence of an infinite variance, was later re-evaluated in the context of the GARCH family of models.

The search for better statistical models which capture all empirical regularities exhibited by speculative price data is still ongoing and the challenge lies in specifying a statistical model that is able to capture the mechanism of speculative price data in a unifying and consistent framework. A promising approach to statistical model specification seems to be the probabilistic reduction approach. In relation to this approach, the comments made by an anonymous discussant on Kendall's 1953 paper are clearly prophetic:

‘the possibility that price changes might be dependent but not correlated, a possibility that might arise in cases of nonlinear relationships or where underlying distributions are nonnormal’. (Cootner, 1964; Introduction, p. 81.)

At the same time it is interesting to re-examine the relationship between the statistical models and certain theoretical issues such as the weak-EMH and major asset pricing models. The EMH can be considered consistent with extensions of the martingale difference process that allow for dynamic heteroskedasticity and non-linear dependence. The results for SP500 and FT-ALL weekly returns provide empirical support for the martingale difference assumption of the EMH. The dynamic volatility of both the US and UK stock market indices can be regarded as forecastable. Nevertheless, one should always bear in mind that market efficiency implies a lot more than a martingale difference returns process (such as rational agents, risk neutrality, linear expected utility, perfect competition) and all we are merely testing is the validity of the statistical mechanism of returns for the stock indices considered. Certain asset pricing models such as the Mean-Variance theorem and the Capital Asset Pricing Model (CAPM) have been generalized to elliptically distributed returns (e.g. Chamberlain, 1985; Owen and Rabinovich, 1985; Ingersoll, 1989). A member of this family of distributions is the Student's  $t$  distribution which seems to be both empirically appropriate (as shown by the review of empirical literature) and theoretically

appropriate since Student's  $t$  distributed returns are completely specified by the mean and variance and is closed under linear transformations.

## Notes

1. This section focuses on parametric dynamic heteroskedastic models. There is also a vast influential literature on non- and semi-parametric models that attempt to capture the stylized facts of speculative prices. Examples of nonparametric model estimation include for instance kernel regression, orthogonal series expansion, artificial neural networks (see Gouriéroux and Monfont, 1992; Pagan and Schwert, 1990, *inter alia*).
2. There is a large number of univariate ARCH parameterizations based on certain modifications of the initial ARCH model, such as the non-linear ARCH, quadratic ARCH, asymmetric ARCH, threshold ARCH, structural ARCH and switching ARCH (see, for instance, Bollerslev *et al.*, 1994). Moreover, (G)ARCH in mean models specify expected returns as a function of conditional volatility. The paper does not make an attempt to evaluate empirically every single ARCH model. Instead the focus of the paper is in dealing with the core of these models (i.e. the ARCH and GARCH parameterizations).

## References

- Akgiray, V. and Booth, G. G. (1988) The stable-law model of stock returns. *Journal of Business and Economic Statistics*, 6, 51–57.
- Alexander, S. (1961) Price movements in speculative markets: Trends and random walks. In P. Cootner (ed.), *The Random Character of Stock Market Prices*. Cambridge, MIT Press, 1964.
- Bachelier, L. (1900) Theory of speculation. In P. Cootner (ed.), *The Random Character of Stock Market Prices*. Cambridge, MIT Press, 1964.
- Bachelier, L. (1914) *Le Jeu, la Chance, et le Hasard*. Flammarion, Paris.
- Baillie, R. (1993) Comment on 'Modeling asset returns with alternative stable distributions'. *Econometric Reviews*, 12, 343–45.
- Baillie, R. (1996) Long memory processes and fractional integration in econometrics. *Journal of Econometrics*, 73, 5–59.
- Ball, C. and Tóuross, W. (1983) Bond price dynamics and options. *Journal of Financial Quantitative Analysis*, 18, 517–31.
- Black, F. (1976) Studies in stock price volatility changes. Proceedings of the 1976 Business Meeting of the Business and Economic Statistics Section, American Statistical Association, 177–181.
- Black, F. and Scholes, M. (1973) The pricing of options and corporate liabilities. *Journal of Political Economy*, 81, 637–659.
- Blattberg, R. and Gonedes, N. (1974) A comparison of Stable and Student's  $t$  distributions as statistical models for stock prices. *Journal of Business*, 47, 244–80.
- Bollerslev, T. (1986) Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31, 307–327.
- Bollerslev, T. (1987) A conditionally heteroskedastic time series model for speculative prices and rates of return. *Review of Economics and Statistics*, 69, 542–47.
- Bollerslev, T., Chou, R. and Kroner, K. (1992) ARCH modeling in finance: A review of the theory and empirical evidence. *Journal of Econometrics*, 52, 5–59.
- Bollerslev, T., Engle, R. F. and Nelson, D. B. (1994) ARCH Models. *Handbook of Econometrics*, IV, Ch. 49, 2659–3038.

- Bollerslev, T. and Mikkelsen, H. O. (1996) Modeling and pricing long memory in stock market volatility. *Journal of Econometrics*, 73, 151–84.
- Breeden, D. (1979) An intertemporal asset pricing model with stochastic consumption and investment opportunities. *Journal of Financial Economics*, 7, 265–96.
- Brock, W. A., Hsieh, D. and LeBaron, B. (1991) *Nonlinear Dynamics, Chaos, and Instability: Statistical Theory and Economic Evidence*. Cambridge, The MIT Press, MA.
- Brock, W. A., Dechert, W. D., Scheinkman J. A. and LeBaron, B. (1996) A test for independence based on the correlation dimension. *Econometric Reviews*, 15, 197–235.
- Cai, J. (1994) A Markov model of switching-regime ARCH. *Journal of Business and Economic Statistics*, 12, 309–16.
- Campbell J. Y., Lo, A. W. and MacKinlay, A. C. (1997) *The Econometrics of Financial Markets*. Princeton, New Jersey, Princeton University Press.
- Chamberlain, G. (1985) A characterization of the distributions that imply mean-variance utility functions. *Journal Economic Theory*, 29, 185–201.
- Chesney, M. and Scott, L. (1989) Pricing European currency options: A comparison of the modified Black-Scholes model and a random variance model. 24, 267–284.
- Cootner, P. H. (1962) Stock prices: random versus systematic changes. *Industrial Management Review*, vol. 3, no. 2, 24–45.
- Cootner, P. H. (1964) *The Random Character of Stock Market Prices*. Cambridge Mass.: MIT Press.
- Cowles, A. (1960) A revision of previous conclusions regarding stock price behavior. *Econometrica*, 28, 909–915.
- Cowles, A. and Jones, H. (1937) Some a posteriori probabilities in stock market action. *Econometrica*, 5, 280–294.
- Cox J., Ingersoll, J. and Ross, S. (1985) An intertemporal general equilibrium model of asset prices. *Econometrica*, 53, 363–84.
- D'Agostino, R. and Stephens, M. A. (1986) *Goodness of Fit Techniques*. New York: Marcel Dekker.
- Diebold, F. (1993) Comment on 'Modeling asset returns with alternative stable distributions'. *Econometric Reviews*, 12, 339–42.
- Ding, Z., Granger, C. W. J. and Engle, R. F. (1993) A long memory property of stock returns and a new model. *Journal of Empirical Finance*, 1, 83–106.
- Ding, Z. and Granger, C. W. J. (1996) Modeling volatility persistence of speculative returns: A new approach. *Journal of Econometrics*, 73, 185–215.
- Doob, J. L. (1953) *Stochastic Processes*. John Wiley and Sons, New York.
- DuMouchel, W. (1973) Stable distributions in statistical inference: 1. Symmetric stable distribution compared to other symmetric long-tailed distributions. *Journal of the American Statistical Association*, 68, 469–477.
- Engle, R. F. (1982) Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50, 987–1008.
- Engle, R. and Bollerslev, T. (1986a) Modeling the persistence of conditional variances. *Econometric Reviews*, 5, 1–50.
- Engle, R. and Bollerslev, T. (1986b) A Reply on 'Modeling the persistence of conditional variances'. *Econometric Reviews*, 5, 81–87.
- Epstein, L. G. and Zin, S. E. (1991) Substitution, risk aversion and the temporal behavior of consumption and asset returns: An empirical analysis. *Journal of Political Economy*, 99, 263–286.
- Fama, E. F. (1965) The behavior 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.
- French, K., Schwert, W. and Stambaugh, R. (1987) Expected stock returns and volatility. *Journal of Financial Economics*, 19, 3–29.
- Ghysels, E., Harvey, A. C. and Renault, E. (1996) Stochastic Volatility. *Handbook of Statistics*, Vol. 14, 119–91.

- Gourieroux, C. and Monfort, A. (1992) Qualitative threshold ARCH models. *Journal of Econometrics*, 52, 159–199.
- Granger, C. W. J. (1981) Some properties of time series data and their use in econometrics model specification. *Journal of Econometrics*, 16, 121–130.
- Granger, C. and Ding, Z. (1994) Stylized facts on the temporal dependence and distributional properties of daily data from speculative prices. Working Paper, University of San Diego.
- Granger, C. W. J. and Joyeux, R. (1980) An introduction to long memory time series models and fractional differencing. *Journal of Time Series Analysis*, 1, 15–39.
- Granger, C. W. J. and Morgestern, O. (1970) *Predictability of Stock Market Prices*. Heatly Lexington, Lexington, MA.
- Hamilton, J. D. (1990) Analysis of time series subject to changes in regime. *Journal of Econometrics*, 45, 39–70.
- Hamilton, J. D. and Susmel, R. (1994) Autoregressive conditional heteroskedasticity and changes in regime. *Journal of Econometrics*, 64, 307–33.
- Hansen, B. (1994) Autoregressive Conditional Density Estimation. *International Economic Review*, 35, 705–730.
- Hagerman, R. (1978) More evidence on the distribution of security returns. *Journal of Finance*, 33, 1213–1220.
- Harrison, J. and Kreps, D. (1979) Martingales and arbitrage in multiperiod security markets. *Journal of Economic Theory*, 20, 381–408.
- Hosking, J. R. M. (1981) Fractional differencing. *Biometrika*, 68, 165–176.
- Hurst, H. E. (1951) Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, 1, 519–43.
- Ingersoll, J. (1989) *Theory of Financial Decision Making*. Rowman & Littlefield, Ch. 4, Appendix B.
- Jarrow, R. and Rosenfeld, E. (1984) Jumping risk and the intertemporal capital asset pricing model. *Journal of Business*, 57, 337–351.
- Kendall, S. (1953) The analysis of economic time series. In Part I: Prices' in Cootner, (ed.) (1964) *The Random Character of Stock Market Prices*, Cambridge Mass.: MIT Press.
- Keynes, J. M. (1936) General theory of employment, interest and money. New York: Harcourt, Brace and World, 1936, 147–167.
- Lamoureux C. G. and Lastrapes, W. D. (1990) Persistence in variance, structural change and the GARCH model. *Journal of Business and Economic Statistics*, 8, 225–34.
- Lau, A., Lau, H. and Wingender, J. (1990) The distribution of stock returns: New evidence against the stable model. *Journal of Business and Economic Statistics*, 8, 217–23.
- LeRoy, S. F. (1989) Efficient capital markets and martingales. *Journal of Economic Literature*, XXVII, 1583–1621.
- Levy, P. (1925) *Calcul des Probabilites*. Gauthier-Villars, Paris.
- Lintner, J. (1965) The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics*, 47, 13–37.
- Lucas, R. E. (1978) Asset prices in an exchange economy. *Econometrica*, 46, 1429–1446.
- Ljung, G. M. and Box, G. E. P. (1979) The likelihood functions of stationary autoregressive-moving average models. *Biometrika*, 66, pp. 265–270.
- 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, 421–40.
- Mandelbrot, B. (1966) Forecasts of future prices, unbiased markets and the 'martingale' models. *Journal of Business*, 39, 1, 242–255.
- Mandelbrot, B. (1967) The variation of some other speculative prices. *Journal of Business*, 36, 394–419.
- Mandelbrot, B. (1969) Long-run linearity, locally Gaussian process, H-spectra and infinite variances. *International Economic Review*, 10, 82–111.



- Mandelbrot, B. (1972) Statistical methodology of non periodic cycles: From the covariance to the R/S analysis. *Annals of Economic and Social Measurement*, 1, 259–290.
- Markowitz, H. (1952) Portfolio Selection. *Journal of Finance*, 7, 77–91.
- McCulloch, J. H. (1997) Measuring Tail Thickness to Estimate the Stable Index  $\alpha$ : A Critique. 15, 74–81.
- McGuirk A., Robertson, J. and Spanos, A. (1993) Modeling exchange rate dynamics: Thick-tails and non-linear dependence. *Econometric Reviews*, 12, 33–36.
- McLeod, A. I. and Li, W. K. (1983) Diagnostic checking ARMA time series models using squared-residual autocorrelations. *Journal of Time Series Analysis*, 4, 239–273.
- Merton, R. (1973) An intertemporal capital asset pricing model *Econometrica*, 41, 867–887.
- Merton, R. C. (1990) *Continuous-Time Finance*. Basil Blackwell.
- Mills, T. C. (1995) Modelling skewness and kurtosis in the London Stock Exchange FT-SE index return distributions. *The Statistician*, 44, 323–32.
- Mittnik, S. and Rachev, S. T. (1993a) Modeling asset returns with alternative stable paretian distributions. *Econometric Reviews*, 12, 261–330.
- Mittnik, S. and Rachev, S. T. (1993b) Reply to comments on ‘Modeling asset returns with alternative stable distributions’ and some extentions. *Econometric Reviews*, 12, 347–389.
- Moore, A. B. (1962) Some characteristics of changes in common stock prices. In P. Cootner (ed.), (1964) *The Random Character of Stock Market Prices*. Cambridge Mass.: MIT Press.
- Mossin, J. (1966) Equilibrium in a capital asset market. *Econometrica*, 35, 768–783.
- Nelson, D. B. (1990) Stationarity and persistence in the GARCH(1, 1) model. *Econometric Theory*, 6, 318–334.
- Nelson, D. B. (1991) Conditional heteroskedasticity in asset prices: A new approach. *Econometrica*, 59, 347–370.
- Officer, R. (1972) The distribution of stock returns. *Journal of the American Statistical Association*, 67, 807–812.
- Osborne, M. (1959) Brownian motopn in stock market. *Operational Research*, 7, 145–173.
- Osborne, M. (1962) Periodic structure in the brownian motion of stock prices. In P. Cootner (ed.), (1964) *The Random Character of Stock Market Prices*. Cambridge Mass.: MIT Press.
- Owen, J. and Rabinovich, R. (1983) On the class of elliptical distributions and their applications to the theory of portfolio choice. *Journal of Finance*, 38, 745–52.
- Pagan, A. R. (1996) The econometrics of financial markets. *Journal of Empirical Finance*, 3, 15–102.
- Pagan, A. R. and Schwert, G. W. (1990) Alternative models for conditional stock volatility. *Journal of Econometrics*, 45, 267–290.
- Phillips, P. (1993) Comment on ‘Modeling asset returns with alternative stable distributions’. *Econometric Reviews*, 12, 331–338.
- Phillips, P. and Loretan, M. (1994) Testing covariance stationarity under moment condition failure with an application to common stock returns. *Journal of Empirical Finance*.
- Praetz, P. (1972) The distribution of share price changes. *Journal of Business*, 45, 49–55.
- Roberts, H. V. (1959) Stock market ‘patterns’ and financial analysis: methodological suggestions. *The Journal of Finance*, XIV, 1.
- Samuelson, P. (1965) Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review*, 6, 41–49.
- Sawyer, K. R. (1993) Continuous time financial models: statistical applications of stochastic processes. In G. S. Maddala, C. R. Rao and H. D. Vinod, (eds), *Handbook of Statistics*, Vol. 0.11.
- Schaefer, S. and Schwartz, E. (1987) Time dependent variance and the pricing of bond options. *Journal of Finance*, 42, 1113–1128.

- Schwert, W. (1989) Why does stock market volatility change over time? *Journal of Finance*, 44, 1115–1153.
- Sharpe, W. (1964) Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, 19, 425–442.
- Silverman, B. W. (1986) *Density Estimation for Statistics and Data Analysis*. London, Chapman Hall.
- Slutsky, E. E. (1937) The summation of random causes as the source of cyclic processes. *Econometrica*, V, 105–146.
- Spanos, A. (1986) *Statistical Foundations of Econometric Modelling*. Cambridge University Press, Cambridge.
- Spanos, A. (1991) The Student's  $t$  autoregressive model with conditional heteroskedasticity. Working Paper, VPI & University of Cyprus.
- Spanos, A. (1994) On modeling heteroskedasticity: the Student's  $t$  and elliptical linear regression models. *Econometric Theory*, 10, 286–315.
- Spanos, A. (1995) On Theory Testing in Econometrics: Modeling with nonexperimental data. *Journal of Econometrics*, 67, 189–226.
- Spanos, A. (1999) *An Introduction to Modern Econometrics*. Cambridge University Press, Cambridge.
- Steiger, T. (1964) A test of nonrandomness in stock price changes. In P. H. Cootner, (ed.), *The Random Character of Stock Market Prices*. (1964) Cambridge Mass.: MIT Press.
- Tucker, A. L. (1992) A reexamination of finite and infinite-variance distributions as models of daily stock returns. *Journal of Business and Economic Statistics*, 10, 73–81.
- Vasicek, O. (1977) An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5, 177–188.
- Working, H. (1934) A random difference series for use in the analysis of time series. *American Statistical Association Journal*, 29, 11–24.
- Working, H. (1960) Note on the correlation of first differences of averages in a random chain. *Econometrica*, 28, 916–918.