ARCH Models for Financial Applications

Evdokia Xekalaki • Stavros Degiannakis

Department of Statistics
Athens University of Economics and Business, Greece

WILEY
A John Wiley and Sons, Ltd, Publication

ARCH Models for Financial Applications

ARCH Models for Financial Applications

Evdokia Xekalaki • Stavros Degiannakis

Department of Statistics
Athens University of Economics and Business, Greece

WILEY
A John Wiley and Sons, Ltd, Publication

This edition first published 2010

Registered office
John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com.

Library of Congress Cataloguing-in-Publication Data

Xekalaki, Evdokia.
ARCH models for financial applications / Evdokia Xekalaki, Stavros Degiannakis.
p. cm.
Includes bibliographical references and index.
ISBN 978-0-470-06630-0 (cloth)
1. Finance–Mathematical models. 2. Autoregression (Statistics) I. Degiannakis, Stavros. II. Title.
HG106.X45 2010
332.01'519536–dc22

2009052104

A catalogue record for this book is available from the British Library.

ISBN 978-0-470-06630-0 (H/B)

Set in 10/12pt Times by Thomson Digital, Noida, India
Printed and bound in Great Britain by TJ International, Padstow, Cornwall

To my husband and my son, a wonderful family

Evdokia Xekalaki

To the memory of the most important person in my life, my father Antonis, and to my mother and my brother

Stavros Degiannakis

Contents

Preface

There has been wide interest throughout the financial literature on theoretical and applied problems in the context of ARCH modelling. While a plethora of articles exists in various international journals, the literature has been rather sparse when it comes to books with an exclusive focus on ARCH models. As a result, students, academics in the area of finance and economics, and professional economists with only a superficial grounding in the theoretical aspects of econometric modelling, while able to understand the basic theories about model construction, estimation and forecasting, often fail to get a grasp of how these can be used in practice.

The present book addresses precisely these issues by interweaving practical questions with approaches hinging on financial and statistical theory: we have adopted an interactional exposition of the ARCH theory and its implementation throughout. This is a book of practical orientation and applied nature intended for readers with a basic knowledge of time series analysis wishing to gain an aptitude in the applications of financial econometric modelling. Balancing statistical methodology and structural descriptive modelling, it aims to introduce readers to the area of discrete time applied stochastic volatility models and to help them acquire the ability to deal with applied economic problems. It provides background on the theory of ARCH models, but with a focus on practical implementation via applications to real data (the accompanying CD-ROM provides programs and data) and via examples worked with econometrics packages (EViews and the G@RCH module for the Ox package) with step-by-step explanations of their use. Readers are familiarized with theoretical issues of ARCH models from model construction, fitting and forecasting through to model evaluation and selection, and will gain facility in employing these models in the context of financial applications: volatility forecasting, value-at-risk forecasting, expected shortfall estimation, and volatility forecasts for pricing options.

Chapter 1 introduces the concept of an autoregressive conditionally heteroscedastic (ARCH) process and discusses the effects that various factors have on financial time series such as the leverage effect, the non-trading period effect, and the non-synchronous trading effect. Chapter 2 provides an anthology of representations of ARCH models that have been considered in the literature. Estimation and simulation of the models is discussed, and several misspecification tests are provided. Chapter 3 deals with fractionally integrated ARCH models and discusses a series of tests for testing the hypothesis of normality of the standardized residuals. Chapter 4 familiarizes readers with the use of EViews in obtaining volatility forecasts. Chapter 5 treats the case of ARCH models with non-normally distributed standardized innovations – in particular, models with innovations with Student *t*, beta, Paretian or Gram–Charlier type distributions, as well as generalized error distributions. Chapter 6 acquaints readers with the use of G@RCH in volatility forecasting. Chapter 7 introduces realized volatility as an alternative volatility measure. The use of high-frequency returns to compute volatility at a lower frequency and the prediction of volatility with ARFIMAX models are presented. Chapter 8 illustrates applications of volatility forecasting in risk management and options pricing. Step-by-step empirical applications provide an insight into obtaining value-at-risk estimates and expected shortfall forecasts. An options trading game driven by volatility forecasts produced by various methods of ARCH model selection is illustrated, and option pricing models for asset returns that conform to an ARCH process are discussed. Chapter 9 introduces the notion of implied volatility and discusses implied volatility indices and their use in ARCH modelling. It also discusses techniques for forecasting implied volatility. Chapter 10 deals with evaluation and selection of ARCH models for forecasting applications. The topics of consistent ranking and of proxy measures for the actual variance are extensively discussed and illustrated via simulated examples. Statistical tests for testing whether a model yields statistically significantly more accurate volatility forecasts than its competitors are presented, and several examples illustrating methods of model selection are given. Finally, Chapter 11 introduces multivariate extensions of ARCH models and illustrates their estimation using EViews and G@RCH.

The contents of the book have evolved from lectures given to postgraduate and final-year undergraduate students at the Athens University of Economics and Business and at the University of Central Greece. Readers are not expected to have prior knowledge of ARCH models and financial markets; they only need a basic knowledge of time series analysis or econometrics, along with some exposure to basic statistical topics such as inference and regression, at undergraduate level.

The book is primarily intended as a text for postgraduate and final-year undergraduate students of economic, financial, business and statistics courses. It is also intended as a reference book for academics and researchers in applied statistics and econometrics, and doctoral students dealing with volatility forecasting, risk evaluation, option pricing, model selection methods and predictability. It can also serve as a handbook for consultants as well as traders, financial market practitioners and professional economists wishing to gain up-to-date expertise in practical issues of financial econometric modelling. Finally, graduate students on master's courses holding degrees from different disciplines may also benefit from the practical orientation and applied nature of the book.

Writing this book has been both exciting and perplexing. We have tried to compile notions, theories and practical issues that by nature lie in areas that are intrinsically complex and ambiguous such as those of financial applications. From numerous possible topics, we chose to include those that we judged most essential. For each of these, we have provided an extensive bibliography for the reader wishing to go beyond the material covered in the book.

We would like to extend our thanks to the several classes of students whose queries and comments in the course of the preparation of the book helped in planning our approach to different issues, in deciding the depth to which chosen topics should be covered and in clearing up ambiguities. We are also grateful to the many people (university colleagues, researchers, traders and financial market practitioners), who by occasional informal exchange of views have had an influence on these aspects as well. Much of the material in this book was developed while the first author was on sabbatical leave at the Department of Statistics, University of California, Berkeley, which she gratefully acknowledges for providing her with a wonderful work environment. Moreover, it has been a pleasure to work on this project with Kathryn Sharples, Simon Lightfoot, Susan Barclay, Richard Davies, Heather Kay and Ilaria Meliconi at John Wiley & Sons, Ltd, copy editor Richard Leigh who so attentively read the manuscript and made useful editing suggestions, and Poirei Sanasam at Thomson Digital. Most importantly, we wish to express our gratitude to our families for their support.

May 2009

Evdokia Xekalaki
Stavros Degiannakis

Notation

$\circ$	Hadamard (elementwise) product
$A(L)$	Polynomial of ARCH
AD	Anderson–Darling statistic
AIC	Akaike information criterion
$\mathbf{B}$	Matrix of unknown parameters in a multivariate regression model.
$B(.,.)$	Cumulative distribution function of the binomial distribution
$B(L)$	Polynomial of GARCH
$B(t)$	Standard Brownian motion
$c(.)$	Smooth function on [0,1].
c_i	Autoregressive coefficients
$C(L)$	Polynomial of AR
$\mathbf{C}_t$	Matrix of conditional correlations
$C_t^{(\tau)}$	Call option at time t, with τ days to maturity
$C_{t+1\|t}^{(\tau)}$	Call option at time $t+1$ given the information available at time t, with τ days to maturity
CGR	Correlated gamma ratio distribution
CM	Cramér–von Mises statistic
d	Exponent of the fractional differencing operator $(1-L)^d$ in FIGARCHmodels
$\tilde{d}$	Exponent of the fractional differencing operator$(1-L)\tilde{d}$ in ARFIMAX models
$\tilde{\tilde{d}}$	Integer differencing operator
d_i	Moving average coefficients
dt_i	Duration or interval between two transactions, $dt_i \equiv t_i - t_{i-1}$
$D(L)$	Polynomial of MA
$DM_{(A,B)}$	Diebold–Mariano statistic
$ES_{t+\tau\|t}^{(1-p)}$	Expected shortfall τ days ahead.
$ES_{t+1\|t}^{(1-p)}$	Expected shortfall forecast at time $t+1$ based on information available at time t, at $(1-p)$ probability level.
$f(.)$	Probability density function
$f_a(.)$	a-quantile of the distribution with density function $f(.)$
$f_{(BG)}(x, y; \widehat{T}, \rho)$	Probability density function of the bivariate gamma distribution

$f_{(CGR)}\left(x; 2^{-1}\ \widehat{T}, \rho\right)$	Probability density function of the correlated gamma ratio distribution
$f_d(0)$	Spectral density at frequency zero
$f_{(GC)}(z_t; v, g)$	Probability density function of Gram–Charlier distribution
$f_{(GED)}(z_t; v)$	Probability density function of generalized error distribution
$f_{(GT)}(z_t; v, g)$	Probability density function of generalized t distribution
$f_{(SGED)}(z_t; v, \theta)$	Probability density function of skewed generalized error distribution
$f_{(skT)}(z_t; v, g)$	Probability density function of skewed Student t distribution
$f_{(t)}(z_t; v)$	Probability density function of standardized Student t distribution.
$F(.)$	Cumulative distribution function
${}_2F_2(.,.;.,.;.)$	Generalized hypergeometric function
F_{il}	Filter
$FoEn_{(1,2)}$	Forecast encompassing statistic of Harvey *et al.* (1998)
$g(.)$	Functional form of conditional variance
g_{s_t}	Constant multiplier in SWARCH model
GED	Generalized error distribution
GT	Generalized t distribution
$\mathbf{H}_t$	Conditional covariance matrix of multivariate stochastic process, $\mathbf{H}_t \equiv V_{t-1}(\mathbf{y}_t)$
Hit_t	Index of VaR violation minus expected ratio of violations, $Hit_t = \tilde{I}_t - p$
HQ	Hannan and Quinn information criterion
$\mathbf{i}$	Vector of ones.
I_t	Information set
$\tilde{I}_t$	Index of VaR violations
JB	Jarque–Bera test
k	Dimension of vector of unknown parameters β
$\tilde{k}$	Number of slices
K	Exercise (or strike) price
$\tilde{K}$	Number of regimes in SWARCH model
$\dddot{K}$	Number of factors in the $\dddot{K}$ Factor ARCH(p, q) model
KS	Kolmogorov–Smirnov statistic
KS^*	Kuiper statistic
Ku	Kurtosis
l	Order of the moving average model
l_i	Eigenvalue
$l_t(.)$	Log-likelihood function for the tth observation
log	Natural logarithm
L	Lag operator
$L_T(.)$	Full-sample log-likelihood function based on a sample of size T
LR_{cc}	Christoffersen's likelihood ratio statistic for conditional coverage.

LR_{in}	Christoffersen's likelihood ratio statistic for independence of violations
LR_{un}	Kupiec's likelihood ratio statistic for unconditional coverage.
m	Number of intraday observations per day
$mDM_{(A,B)}$	Modified Diebold–Mariano statistic
$MGN_{(A,B)}$	Morgan–Granger–Newbold statistic
n	Dimension of the multivariate stochastic process, $\{\mathbf{y}_t\}$
n_{ij}	Number of points in time with value i followed by j
N	Total number of VaR violations, $N = \Sigma_{t=1}^{\tilde{T}} \tilde{I}_t$
$N(.)$	Cumulative distribution function of the standard normal distribution
NRT_t	Net rate of return at time t from trading an option.
p	Order of GARCH form
$p_{i,t}$	Filtered probability in regime switch models (the probability of the market being in regime i at time t)
p_t	Switching probability in regime switch models
$P(.)$	Probability
$P_t^{(\tau)}$	Put option at time t, with τ days to maturity
$P_{t+1\|t}^{(\tau)}$	Put option at time $t+1$ given the information available at time t, with τ days to maturity
$\tilde{q}$	Order of BEKK$(p, q, \tilde{q})$ model.
q_t	Switching probability in regime switch models
rf_t	Rate of return on a riskless asset
R^2	Coefficient of multiple determination
RT_t	Rate of return at time t from trading an option
s_t	Regime in SWARCH model.
$\breve{y}_{it}^*$	Risk-neutral log-returns at time t.
q	Order of ARCH form
$Q^{(LB)}$	Ljung–Box statistic
r_j	Autocorrelation of squared standardized residuals at j lags
skT	Skewed Student t distribution
S_t	Market closing price of asset at time t
$\tilde{S}(T)$	Terminal stock price adjusted for risk neutrality
SBC	Schwarz information criterion
$SGED$	Skewed generalized error distribution
SH	Shibata information criterion
Sk	Skewness
$SPA_{(i^*)}$	Superior predictive ability statistic for the benchmark model i^*
T	Number of total observations, $T = \tilde{T} + \breve{T}$
$\tilde{T}$	Number of observations for out-of-sample forecasting
$\breve{T}$	Number of observations for rolling sample
$\widehat{T}$	Number of observations for model selection methods in out-of-sample evaluation $v_t = \varepsilon_t^2 - \sigma_t^2$
$VaR_t^{(1-p)}$	VaR at $1-p$ probability level at time t

$VaR^{(1-p)}_{t+1\|t}$	VaR forecast at time $t+1$ based on information available at time t, at $(1-p)$ probability level
$VaR^{(1-p)}_{t+\tau\|t}$	VaR τ days ahead
$\widehat{VaR}^{(1-p)}_{t}$	VaR in-sample estimate at time t, at $1-p$ probability level
$vech(.)$	Operator stacks the columns of square matrix.
w	Vector of estimated parameters for the density function f
$\breve{w}$	Number of parameters of vector w
w_i	Weight
X	Transaction cost
y_t	Log-returns
$\mathbf{y}_t$	Multivariate stochastic process
$y_{t+1\|t(i)}$	One-step-ahead conditional mean at time $t+1$ based on information available at time t, from model i
$y^{(BC)}_t$	Box–Cox transformed variable of y_t
$z_{t+1\|t}$	One-step-ahead standardized prediction errors at time $t+1$ based on information available at time t.
α_c	Measure of cost of capital opportunity
β	Vector of unknown parameters in a regression model
$\gamma_d(i)$	Sample autocovariance of ith order
γ_t	Dividend yield
ε_t	Innovation process for the conditional mean of multivariate stochastic process, $\varepsilon_t \equiv \mathbf{y}_t - \boldsymbol{\mu}_t$
$\varepsilon_{t+1\|t}$	One-step-ahead standardized prediction errors at time $t+1$ based on information available at time t
$\tilde{\varepsilon}_t$	Innovation process in SWARCH model
θ	Vector of estimated parameters for the conditional mean and variance
$\theta^{(t)}_{(i)}$	Vector of estimated parameters for the conditional mean and variance at time t from model i
$\breve{\theta}$	Number of parameters of vector θ
κ	Order of the autoregressive model
μ	Instantaneous expected rate of return
$\mu(.)$	Functional form of conditional mean
μ_t	Predictable component of conditional mean
$\boldsymbol{\mu}_t$	Conditional mean of multivariate stochastic process, $\boldsymbol{\mu}_t \equiv E_{t-1}(\mathbf{y}_t)$
v	Tail-thickness parameter
π_{ij}	Percentage of points in time with value i followed by j, $\pi_{ij} = n_{ij}/\Sigma_j n_{ij}$
ρ	Correlation coefficient
σ	Instantaneous variance of the rate of return
$\sigma_{i,j,t}$	Conditional covariance between asset returns i and j at time t.
σ^2_{t+1}	True, but unobservable, value of the variance at time $t+1$

$\sigma^2_{t+1\|t}$	One-step-ahead conditional variance at time $t+1$ based on information available at time t.
$\sigma^2_{t+1\|t(i)}$	One-step-ahead conditional variance at time $t+1$ based on information available at time t, from model i.
$\sigma^2_{t+s\|t}$	s-days-ahead conditional variance at time $t+s$ based on information available at time t
$\sigma^{2(\tau)}_{t+1}$	True, but unobservable, value of variance for a period of τ days, from $t+1$ until $t+\tau$
$\sigma^{2(\tau)}_{t+1\|t}$	Variance forecast for a period of τ days, from $t+1$ until $t+\tau$, given the information available at time t
$\tilde{\sigma}^{2(\tau)}_{t+1}$	Proxy for true, but unobservable, value of variance for a period of τ days, from $t+1$ until $t+\tau$
$\bar{\sigma}^{(\tau)}_{t+1\|t}$	Average standard deviation forecasts from $t+1$ up to $t+\tau$, given the information available at time t, $\bar{\sigma}^{(\tau)}_{t+1\|t} = \sqrt{\tau^{-1}\sum_{i=1}^{\tau}\sigma^2_{t+i\|t}} = \sqrt{\tau^{-1}\sigma^{2(\tau)}_{t+1\|t}}$
$\hat{\sigma}^2_{t+1}$	In-sample conditional variance at time $t+1$ based on the entire available data set T
$\hat{\sigma}^{2(\tau)}_{t+1}$	In-sample conditional variance for a period of τ days, from $t+1$ until $t+\tau$, based on the entire available data set T
$\hat{\sigma}^{2(RV)}_{(un),t+1}$	In-sample realized volatility at time $t+1$ based on the entire available data set T
$\sigma^{2(RV)}_{t+1}$	Observable value of the realized variance at time $t+1$
$\sigma^{2(RV)(\tau)}_{t+1}$	Observable value of the realized variance for a period of τ days, from $t+1$ until $t+\tau$
$\sigma^{2(RV)}_{(un),t+1\|t}$	One-day-ahead conditional realized variance at time $t+1$ based on information available at time t
τ	Point in time (i.e. days) for out-of-sample forecasting. Also days to maturity for options
υ_t	Vector of predetermined variables included in I_t
$\phi(.)$	Functional form of conditional variance in conditional mean in GARCH-M model
$\Phi(.;.;.)$	Confluent hypergeometric function
$\Phi(L)$	Polynomial of FIGARCH
$\varphi(t,a,\beta,\sigma,\mu)$	Characteristic function of stable Paretian distribution
$\chi^2_{(g)}$	Pearson's chi-square statistic
ψ	Vector of estimated parameters for the conditional mean, variance and density function, $\psi' = (\theta', w')$
$\psi(.)$	Euler psi function
$\breve{\psi}$	Number of parameters of vector ψ, $\breve{\psi} = \breve{\theta} + \breve{w}$
$\hat{\psi}^{(T)}$	Maximum likelihood estimator of ψ based on a sample of size T

$\bar{\Psi}^{(\tau)}_{(SE)(i)}$	Mean squared error loss function for model i volatility forecasts τ days ahead, i.e. $\bar{\Psi}^{(\tau)}_{(SE)(i)} = \tilde{T}^{-1}\sum_{t=1\tilde{T}}\left(\sigma^{2(\tau)}_{t+1\mid t(i)} - \sigma^{2(\tau)}_{t+1(i)}\right)^2$
$\Psi^{(\tau)}_{(SE)t(i)}$	Squared error loss function for models i volatility forecasts τ days ahead
$\bar{\Psi}^{(\tau)}$	Average of a loss function for τ-days-ahead volatility forecasts, i.e. $\bar{\Psi}^{(\tau)} = \tilde{T}^{-1}\sum_{t=1}^{\tilde{T}}\Psi^{(\tau)}_t$
$\Psi^{(\tau)}_{t(i)}$	Loss function that measures the distance between actual volatility over a τ-day period and model i volatility forecast over the same period
$\Psi_{t(i)}$	Loss function that measures the distance between one-day actual volatility and its forecast by model i, i.e. $\Psi_{t(i)} \equiv \Psi^{(1)}_{t(i)}$
$\Psi^{(\tau)}_{t(i^*,i)}$	Difference of loss functions (loss differential) of models i^* and i, $\Psi^{(\tau)}_{t(i^*,i)} = \Psi^{(\tau)}_{t(i^*)} - \Psi^{(\tau)}_{t(i)}$
$\bar{\Psi}^{(\tau)}_{(i^*,i)}$	Sample mean loss differential of models i^* and i, $\bar{\Psi}^{(\tau)}_{(i^*,i)} = \tilde{T}^{-1}\sum_{t=1}^{\tilde{T}}\Psi^{(\tau)}_{t(i^*,i)}$

1

What is an ARCH process?

1.1 Introduction

Since the first decades of the twentieth century, asset returns have been assumed to form an independently and identically distributed (i.i.d.) random process with zero mean and constant variance. Bachellier (1900) was the first to contribute to the theory of random walk models for the analysis of speculative prices. If $\{P_t\}$ denotes the discrete time asset price process and $\{y_t\}$ the process of continuously compounded returns, defined by $y_t = 100 \log (P_t/P_{t-1})$, the early literature viewed the system that generates the asset price process as a fully unpredictable random walk process:

$$\begin{aligned} P_t &= P_{t-1} + \varepsilon_t \\ \varepsilon_t &\overset{i.i.d.}{\sim} N(0, \sigma^2), \end{aligned} \tag{1.1}$$

where ε_t is a zero-mean i.i.d. normal process. Figures 1.1 and 1.2 show simulated $\{P_t\}_{t=1}^T$ and $\{y_t\}_{t=1}^T$ processes for $T = 5000$, $P_1 = 1000$ and $\sigma^2 = 1$.

However, the assumptions of normality, independence and homoscedasticity do not always hold with real data.

Figures 1.3 and 1.4 show the daily closing prices of the London Financial Times Stock Exchange 100 (FTSE100) index and the Chicago Standard and Poor's 500 Composite (S&P500) index. The data cover the period from 4 April 1988 until 5 April 2005. At first glance, one might say that equation (1.1) could be regarded as the data-generating process of a stock index. The simulated process $\{P_t\}_{t=1}^T$ shares common characteristics with the FTSE100 and the S&P500 indices.[1] As they are clearly

[1] The aim of the visual comparison here is not to ascertain a model that is closest to the realization of the stochastic process (in fact another simulated realization of the process may result in a path quite different from that depicted in Figure 1.1). It is merely intended as a first step towards enhancing the reader's thinking about or conceiving of these notions by translating them into visual images. Higher-order quantities, such as the correlation, absolute correlation and so forth, are much more important tools in the analysis of stochastic process than their paths.

ARCH Models for Financial Applications Evdokia Xekalaki and Stavros Degiannakis

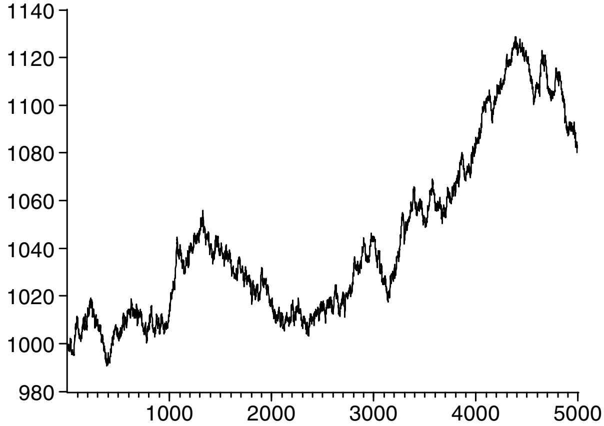

Figure 1.1 *Simulated* $\{P_t\}$ *process, where* $P_t = P_{t-1} + \varepsilon_t$, $P_1 = 1000$ *and* $\varepsilon_t \overset{i.i.d.}{\sim} N(0,1)$.

non-stationary, the autocorrelations presented in Figure 1.5 are marginally less than unity in any lag order. Figure 1.6 plots the distributions of the daily FTSE100 and S&P500 indices as well as the distribution of the simulated process $\{P_t\}_{t=1}^{T}$. The density estimates are based on the normal kernel with bandwidths method calculated according to equation (3.31) of Silverman (1986). S&P500 closing prices and the simulated process $\{P_t\}_{t=1}^{T}$ have similar density functions.

However, this is not the case for the daily returns. Figures 1.7 and 1.8 depict the FTSE100 and S&P500 continuously compounded daily returns, $\{y_{FTSE100,t}\}_{t=1}^{T}$ and

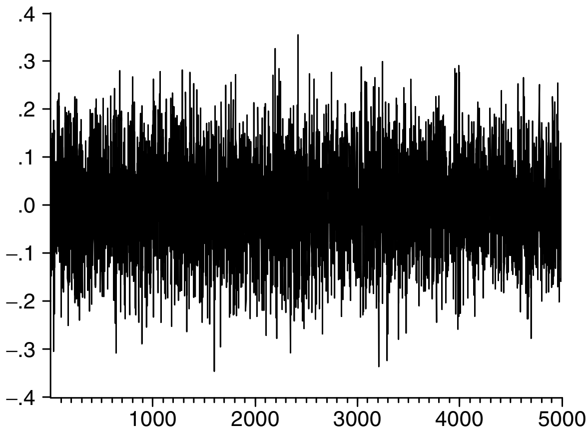

Figure 1.2 *Simulated* $\{y_t\}$ *process, where* $y_t = 100\log(P_t/P_{t-1})$, $P_t = P_{t-1} + \varepsilon_t$, $P_1 = 1000$ *and* $\varepsilon_t \overset{i.i.d.}{\sim} N(0,1)$.

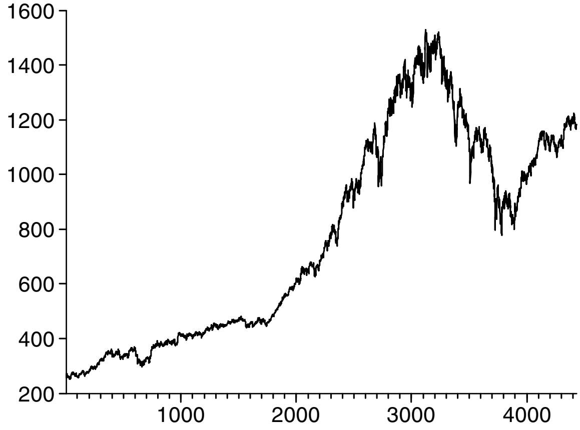

Figure 1.3 S&P500 equity index closing prices from 4 April 1988 to 5 April 2005.

$\{y_{SP500,t}\}_{t=1}^{T}$, while Figure 1.9 presents the autocorrelations of $\{y_t\}_{t=1}^{T}$, $\{y_{FTSE100,t}\}_{t=1}^{T}$ and $\{y_{SP500,t}\}_{t=1}^{T}$ for lags of order 1, ..., 35. The 95% confidence interval for the estimated sample autocorrelation is given by $\pm 1.96/\sqrt{T}$, in the case of a process with independently and identically normally distributed components. The autocorrelations of the FTSE100 and the S&P500 daily returns differ from those of the simulated process. In both cases, more than 5% of the estimated autocorrelations are outside the above 95% confidence interval. Visual inspection of Figures 1.7 and 1.8 shows clearly that the mean is constant, but the variance keeps changing over time, so the return series does not appear to be a sequence of i.i.d. random variables. A characteristic of asset returns, which is noticeable from the figures, is the volatility clustering first noted by

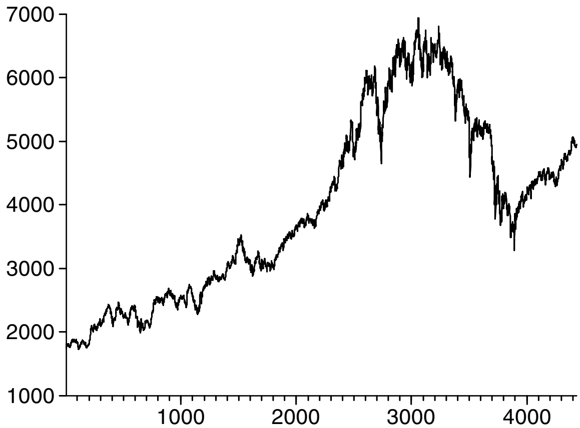

Figure 1.4 FTSE100 equity index closing prices from 4 April 1988 to 5 April 2005.

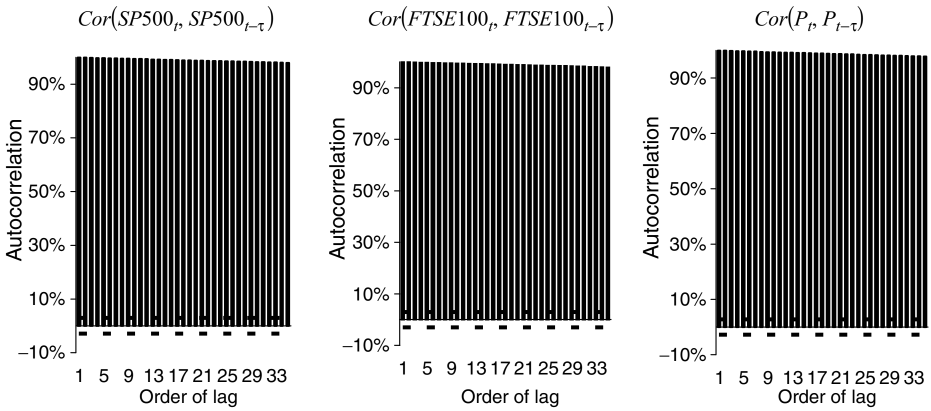

Figure 1.5 Autocorrelation of the S&P500 and the FTSE100 closing prices and of the simulated process $\{P_t\}_{t=1}^{T}$*, for* $\tau = 1(1)35$ *lags. Dashed lines present the 95% confidence interval for the estimated sample autocorrelations given by* $\pm 1.96/\sqrt{T}$*.*

Mandelbrot (1963): 'Large changes tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes'. Fama (1970) also observed the alternation between periods of high and low volatility: 'Large price changes are followed by large price changes, but of unpredictable sign'.

Figure 1.10 presents the histograms of the stock market series. Asset returns are highly peaked (leptokurtic) and slightly asymmetric, a phenomenon observed by Mandelbrot (1963):

> The empirical distributions of price changes are usually too peaked to be relative to samples from Gaussian populations . . . the histograms of price changes are indeed unimodal and their central bells [recall] 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.

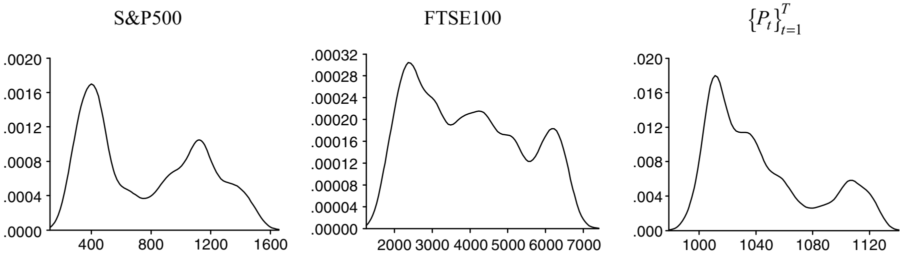

Figure 1.6 Density estimate of the S&P500 and FTSE100 closing prices, and of the simulated process $\{P_t\}_{t=1}^{T}$*.*

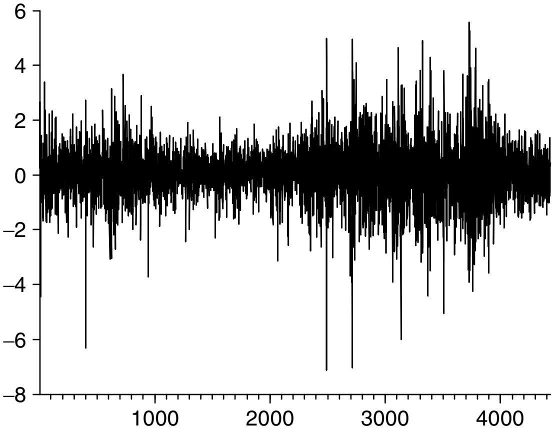

Figure 1.7 S&P500 equity index continuously compounded daily returns from 5 April 1988 to 5 April 2005.

According to Table 1.1, for estimated kurtosis[2] equal to 7.221 (or 6.241) and an estimated skewness[3] equal to -0.162 (or -0.117), the distribution of returns is flat (platykurtic) and has a long left tail relative to the normal distribution. The Jarque and Bera (1980, 1987) test is usually used to test the null hypothesis that the series is normally distributed. The test statistic measures the size of the difference between the skewness, Sk, and kurtosis, Ku, of the series and those of the normal distribution. It is computed as $JB = T\left(Sk^2 + \left((Ku-3)^2/4\right)\right)/6$, where T is the number of observations. Under the null hypothesis of a normal distribution, the JB statistic is χ^2 distributed

[2] Kurtosis is a measure of the degree of peakedness of a distribution of values, defined in terms of a normalized form of its fourth central moment by μ_4/μ_2^2 (it is in fact the expected value of quartic standardized scores) and estimated by

$$Ku = T\sum_{t=1}^{T}(y_t-\bar{y})^4 \Big/ \left(\sum_{t=1}^{T}(y_t-\bar{y})^2\right)^2,$$

where T is the number of observations and $\bar{y}$ is the sample mean, $\bar{y} = \sum_{t=1}^{T} y_t$. The normal distribution has a kurtosis equal to 3 and is called *mesokurtic*. A distribution with a kurtosis greater than 3 has a higher peak and is called *leptokurtic*, while a distribution with a kurtosis less than 3 has a flatter peak and is called *platykurtic*. Some writers talk about *excess kurtosis*, whereby 3 is deducted from the kurtosis so that the normal distribution has an excess kurtosis of 0 (see Alexander, 2008, p. 82).

[3] Skewness is a measure of the degree of asymmetry of a distribution, defined in terms of a normalized form of its third central moment of a distribution by $\mu_3/\mu_2^{3/2}$ (it is in fact the expected value of cubed standardized scores) and estimated by

$$Sk = \sqrt{T}\sum_{t=1}^{T}(y_t-\bar{y})^3 \Big/ \left(\sum_{t=1}^{T}(y_t-\bar{y})^2\right)^{3/2}.$$

The normal distribution has a skewness equal to 0. A distribution with a skewness greater than 0 has a longer right tail is described as *skewed to the right*, while a distribution with a skewness less than 0 has a longer left tail and is described as *skewed to the left*.

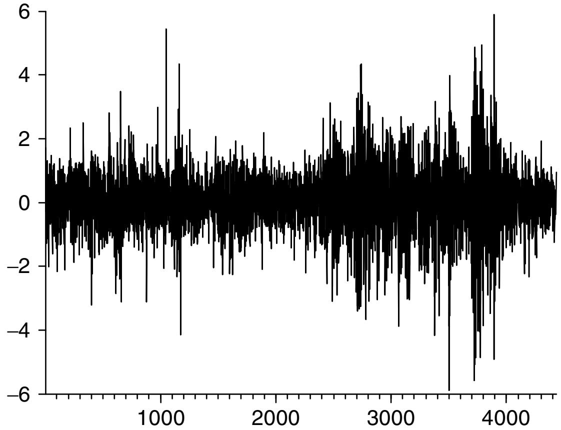

Figure 1.8 FTSE100 equity index continuously compounded daily returns from 5 April 1988 to 5 April 2005.

with 2 degrees of freedom. The two return series were tested for normality using *JB* resulting in a *p*-value that was practically zero, thus signaling non-validity of the hypothesis. Due to the fact that the *JB* statistic frequently rejects the hypothesis of normality, especially in the presence of serially correlated observations, a series of more powerful test statistics (e.g. the Anderson–Darling and the Cramér–von Mises statistics) were also computed with similar results. A detailed discussion of the computation of the aforementioned test statistics is given in Section 3.3.

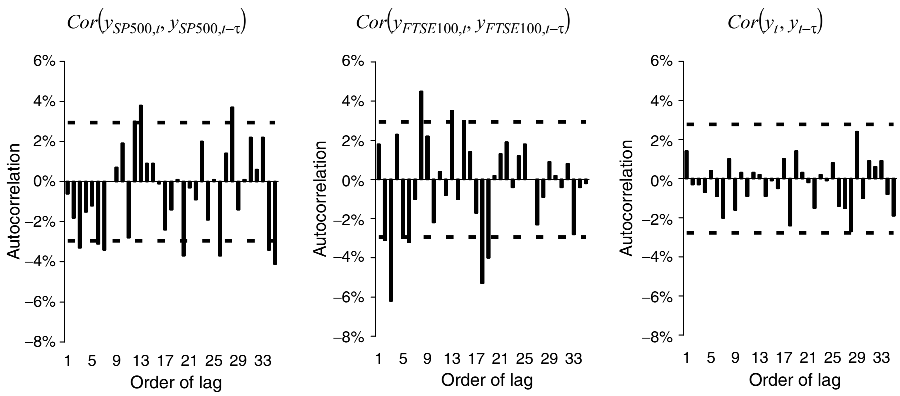

Figure 1.9 Autocorrelation of the S&P500 and FTSE100 continuously compounded daily returns and of the simulated process $\{y_t\}_{t=1}^{T}$, for $\tau = 1(1)35$ lags. Dashed lines present the 95% confidence interval for the estimated sample autocorrelations given by $\pm 1.96/\sqrt{T}$.

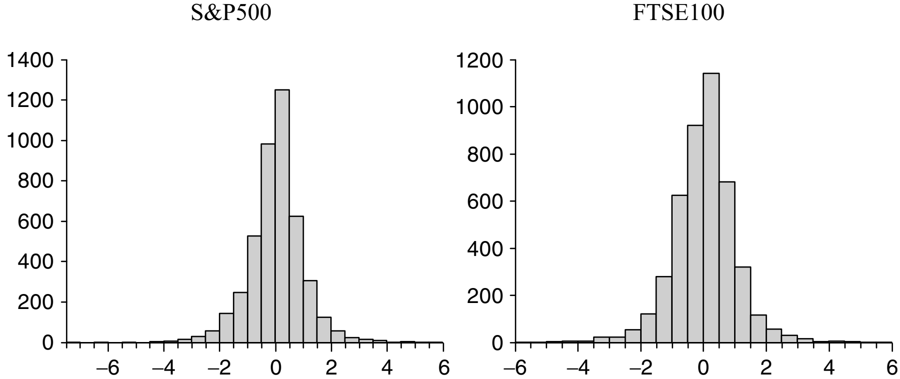

Figure 1.10 Histogram of the S&P500 and FTSE100 index log-returns.

In the 1960s and 1970s, the regularity of leptokurtosis led to a literature on modelling asset returns as i.i.d. random variables having some thick-tailed distribution (Blattberg and Gonedes, 1974; Clark, 1973; Hagerman, 1978; Mandelbrot, 1963, 1964; Officer, 1972; Praetz, 1972). These models, although able to capture the leptokurtosis, could not account for the existence of non-linear temporal dependence such as volatility clustering observed from the data. For example, applying an autoregressive model to remove the linear dependence from an asset returns series and testing the residuals for a higher-order dependence using the Brock–Dechert–Scheinkman (BDS) test (Brock et al., 1987, 1991, 1996), the null hypothesis of i.i.d. residuals was rejected.

Table 1.1 Descriptive statistics of the S&P500 and the FTSE100 equity index returns

	S&P500	FTSE100
Mean	0.034%	0.024%
Standard deviation	15.81%	15.94%
Skewness	−0.162	−0.117
Kurtosis	7.221	6.241
Jarque–Bera	3312.9	1945.6
[*p*-value]	[0.00]	[0.00]
Anderson–Darling	44.3	28.7
[*p*-value]	[0.00]	[0.00]
Cramér–von Mises	8.1	4.6
[*p*-value]	[0.00]	[0.00]

The annualized standard deviation is computed by multiplying the standard deviation of daily returns by $252^{1/2}$, the square root of the number of trading days per year. The Jarque–Bera, Anderson–Darling and Cramér–von Mises statistics test the null hypothesis that the daily returns are normally distributed.

1.2 The autoregressive conditionally heteroscedastic process

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

Before we proceed to the definition of the ARCH model, let us simulate a process able to capture the volatility clustering of asset returns. Assume that the true data-generating process of continuously compounded returns, y_t, has a fully unpredictable conditional mean and a time-varying conditional variance:

$$\begin{aligned} y_t &= \varepsilon_t, \\ \varepsilon_t &= z_t\sqrt{a_0 + a_1\varepsilon_{t-1}^2}, \end{aligned} \tag{1.2}$$

where $z_t \overset{i.i.d.}{\sim} N(0,1)$, z_t is independent of ε_t, $a_0 > 0$ and $0 < a_1 < 1$. The unconditional mean of y_t is $E(y_t) = E(z_t)E\left(\sqrt{a_0 + a_1\varepsilon_{t-1}^2}\right) = 0$, as $E(z_t) = 0$ and z_t and ε_{t-1} are independent of each other. The conditional mean of y_t given the lag values of ε_t is $E(y_t|\varepsilon_{t-1},\ldots,\varepsilon_1) = 0$. The unconditional variance is $V(y_t) = E(\varepsilon_t^2) - E(\varepsilon_t)^2 = E\left(z_t^2\left(a_0 + a_1\varepsilon_{t-1}^2\right)\right) = a_0 + a_1E(\varepsilon_{t-1}^2)$. As $E(\varepsilon_t^2) = E(\varepsilon_{t-1}^2)$, $V(y_t) = a_0(1-a_1)^{-1}$. The conditional variance is $V(y_t|\varepsilon_{t-1},\ldots,\varepsilon_1) = E\left(\varepsilon_t^2|\varepsilon_{t-1},\ldots,\varepsilon_1\right) = a_0 + a_1\varepsilon_{t-1}^2$. The kurtosis of the unconditional distribution equals $E(\varepsilon_t^4)/E(\varepsilon_t^2)^2 = 3(1-a_1^2)/(1-3a_1^2)$. Note that the kurtosis exceeds 3 for $a_1 > 0$ and diverges if a_1 approaches $\sqrt{1/3}$. Figure 1.11 plots the unconditional kurtosis for $0 \le a_1 < 1/\sqrt{3}$.

Both the unconditional and conditional means and the unconditional variance of asset returns remain constant, but the conditional variance has a time-varying character as it depends on the previous values of ε_t. Let us consider equation (1.2) as the true data-generating function and produce 5000 values of y_t. Figure 1.12

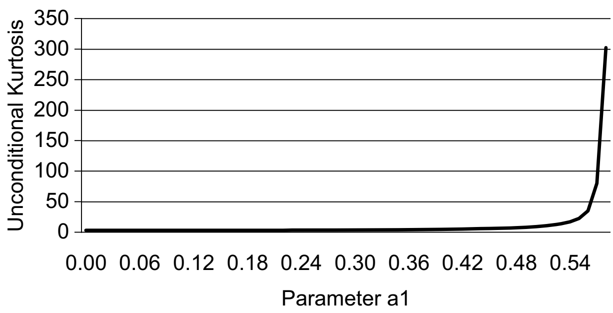

Figure 1.11 *The unconditional kurtosis of* $\varepsilon_t = z_t\sqrt{a_0 + a_1\varepsilon_{t-1}^2}$ *for* $0 \le \alpha_1 < 1/\sqrt{3}$ *and* $a_0 > 0$.