Gilles Teyssière
Alan P. Kirman
Editors

Long Memory in Economics

Springer

Long Memory in Economics

Gilles Teyssière · Alan P. Kirman
Editors

Long Memory in Economics

With 116 Figures and 50 Tables

Springer

Dr. Gilles Teyssière
Université Paris 1
Laboratoire de Statistique Appliquée
et de Modélisation Stochastique (SAMOS)
Panthéon Sorbonne
Centre Pierre Mendès
90 rue de Tolbiac
75634 Paris Cedex 13
France

Professor Alan P. Kirman
Université Aix-Marseille III
L'Ecole des Hautes Etudes
en Sciences Sociales (EHESS)
2 rue de la Charité
13002 Marseille
France

ISBN-10 3-540-22694-X Springer Berlin Heidelberg New York
ISBN-13 978-3-540-22694-9 Springer Berlin Heidelberg New York

Cataloging-in-Publication Data
Library of Congress Control Number: 2005930460

Springer is a part of Springer Science+Business Media

springeronline.com

Printed in Germany

Hardcover-Design: Erich Kirchner, Heidelberg

SPIN 11308669 43/3100-5 4 3 2 1 0 – Printed on acid-free paper

Preface

Long–range dependent, or long–memory, time series are stationary time series displaying a statistically significant dependence between very distant observations. We formalize this dependence by assuming that the autocorrelation function of these stationary series decays very slowly, hyperbolically, as a function of the time lag.

Many economic series display these empirical features: volatility of asset prices returns, future interest rates, etc. There is a huge statistical literature on long–memory processes, some of this research is highly technical, so that it is cited, but often misused in the applied econometrics and empirical economics literature. The first purpose of this book is to present in a formal and pedagogical way some statistical methods for studying long–range dependent processes.

Furthermore, the occurrence of long–memory in economic time series might be a statistical artefact as the hyperbolic decay of the sample autocorrelation function does not necessarily derive from long–range dependent processes. Indeed, the realizations of non-homogeneous processes, e.g., switching regime and change–point processes, display the same empirical features. We thus also present in this book recent statistical methods able to discriminate between the long–memory and change–point alternatives.

Going beyond the purely statistical analysis of economic series, it is of interest to determine which economic mechanisms are generating the strong dependence properties of economic series, whether they are genuine, or spurious. The regularities of the long–memory and change–point properties across economic time series, e.g., common degree of long–range dependence and/or common change–points, suggest the existence of a common economic cause. A promising approach is the use of the class of micro–based models in which the set of economic agents is represented as a (self)–organizing and interacting society whose composition evolves over time, i.e., something resembling the realization of a non–homogeneous stochastic process. Some of these models, inspired by the works of entomologists, are able to mimic some empirical properties of financial and non–financial markets. This implicitly suggests that what is termed as “long–memory” in economics is more complex than a standard (nonlinear) long–range dependent process, and mastering a wide range of statistical tools is a great asset for studying economic time series.

This volume starts with the chapter by Liudas Giraitis, Remis Leipus and Donatas Surgailis, who have been awarded this year the Lithuanian National Prize for Science for their work on “Long–memory: models, limit distributions, statistical inference”. Donatas Surgailis and his numerous former students have made, and are still making, essential contributions to this topic. Their (encyclopedic) survey chapter reviews some recent theoretical findings on ARCH type volatility models. They focus mainly on covariance stationary models which display empirically observed properties which have come to be recognized as ”stylized facts”. One of the major issues to determine is whether the corresponding model for squares r_k^2 of ARCH sequences has long–memory or short memory. It is pointed out that for several ARCH-type models the behavior of $\mathrm{Cov}(r_k^2, r_0^2)$ alone is sufficient to derive the limit distribution of $\sum_k (r_k^2 - Er_k^2)$ and statistical inferences, without imposing any additional (e.g. mixing) assumptions on the dependence structure. This first chapter also discusses ARCH(∞) processes and their modifications such as linear ARCH, bilinear models and stochastic volatility, regime switching stochastic volatility models, random coefficient ARCH and aggregation. They give an overview of the theoretical results on the existence of a stationary solution, dependence structure, limit behavior of partial sums, leverage effect and long–memory property of these models. Statistical estimation of ARCH parameters and testing for change-points are also discussed.

Bhansali, Holland and Kokoszka consider a new and an entirely different approach to modeling phenomena exhibiting long–memory, intermittency and heavy-tailed marginal distributions, namely through the use of chaotic intermittency maps. This class of maps has witnessed considerable development in recent years and it represents an important emerging branch of the subject area of Dynamical Systems Theory. Three principal properties of these maps are relevant and these properties qualify them as a plausible class of models for financial returns. First, unlike some of the standard chaotic maps, the intermittency maps display long–memory and have correlations decaying at a sub-exponential rate, meaning at a polynomial rate or even slower. Secondly, the invariant density of these maps can display ’Pareto’ tails and thus go to zero at a polynomial rate. Thirdly, as their name implies, these maps display intermittency and generate time series, called the orbit of the map, which display intermittent chaos, meaning the orbit of the map alternates between laminar and chaotic regions.

Brousseau analyzes the time series of the euro–dollar exchange rate as the realization of a continuous–time physical process, which implies the use of different degrees of time resolution. The analysis takes into account various statistical indicators, but puts special emphasis on the spectrum of the process. Brousseau finds that this spectrum has an identifiable pattern, which is a core characteristic of the process. Then he simulates a process having the same spectrum, and the behavior of the actual process and of the simulated process are compared using various statistical indicators. It appears that the simulated process provides a good, but not perfect, replication of the behavior of the actual euro–dollar exchange rate.

The next two chapters deal with the issue of change–point detection. Račkauskas and Suquet present the invariance principle by Donsker and Prokhorov, which can be used for analyzing structural changes. They focus on invariance principles with respect to Hölder topologies, as Hölder spaces bring out well variations properties of processes. They present some applications of the Hölderian invariance principles to the problem of testing the stability of a sample against epidemic change–points alternatives.

Lavielle and Teyssière consider the multiple change–point problem for time series, including strongly dependent processes, with an unknown number of change–points. They propose an adaptive method for finding the sequence of change–points $\boldsymbol{\tau}$ with the optimal level of resolution. This optimal segmentation is obtained by minimizing a standard penalized contrast function $J(\boldsymbol{\tau}, \boldsymbol{y}) + \beta \mathrm{pen}(\boldsymbol{\tau})$. The adaptive procedure is such that the optimal segmentation does not strongly depend on the penalization parameter β. This algorithm is applied to the problem of detection of change–points in the volatility of financial time series, and compared with the binary segmentation procedure by Vostrikova.

The chapter by Henry is on the issue of bandwidth selection for semiparametric estimators of the long–memory parameter. The spectral based estimators are derived from the shape of the spectral density at low frequencies, where all but the lowest harmonics of the periodogram are discarded. This allows one to ignore the specification of the short range dynamic structure of the time series, and avoid the bias incurred when the latter is misspecified. Such a procedure entails an order of magnitude loss of efficiency with respect to parametric estimation, but may be warranted when long series (earth scientific or financial) can be obtained. This chapter presents strategies proposed for the choice of bandwidth, i.e., the number of periodogram harmonics used in estimation, with the aim of minimizing this loss of efficiency.

Teyssière and Abry present and study the performance of the semiparametric wavelet estimator for the long–memory parameter devised by Veitch and Abry, and compare this estimator with two semiparametric estimators in the spectral domain: the local Whittle and the "log–periodogram" estimators. The wavelet estimator performs well for a wide range of nonlinear long–memory processes in the conditional mean and the conditional variance, and is reliable for discriminating between change–points and long–range dependence in volatility. The authors also address the issue of the selection of the range of octaves used as regressors by the weighted least squares estimator. It appears that using the feasible optimal bandwidths for either the spectral estimators, surveyed by Henry in the previous chapter, is a useful rule of thumb for selecting the lowest octave. The wavelet estimator is applied to volatility and volume financial time series.

Kateb, Seghier and Teyssière study a fast version of the Levinson–Durbin algorithm, derived from the asymptotic behavior of the first column of the inverse of $T_N(f)$, the $(N+1) \times (N+1)$ Toeplitz matrix with typical element f, the spectral density of a long–memory process. The inversion of $T_N(f)$ with Yule–Walker type equations requires $O(N^3)$ operations, while the Levinson–Durbin algorithm requires $O(N^2)$ elementary operations. In this chapter, an asymptotic behavior of $(T_N(f))$, for large values of N is given so that the computations of the inverse elements are performed in $O(N)$ operations. The numerical results are compared with those given by the Levinson–Durbin algorithm, with particular emphasis on problems of predicting stationary stochastic long–range dependent processes.

Gaunersdorfer and Hommes study a simple nonlinear structural model of endogenous belief heterogeneity. News about fundamentals is an independent and identically distributed random process, but nevertheless volatility clustering occurs as an endogenous phenomenon caused by the interaction between different types of traders, fundamentalists and technical analysts. The belief types are driven by adaptive, evolutionary dynamics according to the success of the prediction strategies as measured by accumulated realized profits, conditioned upon price deviations from the rational expectations fundamental price. Asset prices switch irregularly between two different regimes – periods of small price fluctuations and periods of large price changes triggered by random news and reinforced by technical trading – thus, creating time varying volatility similar to that observed in real financial data.

Cont attempts to model the volatility clustering of asset prices: large changes in prices tend to cluster together, resulting in persistence of the amplitudes of their changes. After recalling various methods for quantifying and modeling this phenomenon, this chapter discusses several economic mechanisms which have been proposed to explain the origin of this volatility clustering in terms of behavior of market participants and the news arrival process. A common feature of these models seems to be a switching between low and high activity regimes with heavy-tailed durations of regimes. Finally, a simple agent-based model is presented, which links such variations in market activity to threshold behavior of market participants and suggests a link between volatility clustering and investor inertia.

Kirmans chapter is devoted to an account of the sort of micro-economic founded models that can give rise to long memory and other stylised facts about financial and economic series. The basic idea is that the market is populated by individuals who have different views about future prices. As time unfolds they may change their ways of forecasting the future. As they do so they change their demands and thus prices. In many of these models the typical rules are "chartist" or extrapolative rules and those based on the idea that prices will revert to some "fundamental" values. In fact, any finite number of rules can be considered. The switching between rules may result in "herding" on some particular rule for a period of time and give rise to long memory and volatility clustering. The first models were used to generate data which was then tested to see if the stylised facts were generated. More recently theoretical results have been obtained which characterise the long run distribution of the price process.

Alfarano and Lux present a very simple model of a financial market with heterogeneous interacting agents capable of reproducing empirical statistical properties of returns. In this model, the traders are divided into two groups, fundamentalists and chartists, and their interactions are based on herding mechanism. The statistical analysis of the simulated data shows long-term dependence in the auto-correlations of squared and absolute returns and hyperbolic decay in the tail of the distribution of the raw returns, both with estimated decay parameters in the same range like empirical data. Theoretical analysis, however, excludes the possibility of "true" scaling behavior because of the Markovian nature of the underlying process and the finite set of possible realized returns.

The purpose of the chapter by de Peretti is to determine whether hysteretic series can be confused with long–memory series. The hysteretic effect is a persistence in the series like the long–memory effect, although hysteretic series are not mean reverting whereas long–memory series are. Hysteresis models have in fact a short memory, since dominant shocks erase the memory of the series, and the persistence is due to permanent and non-reverting state changes at a microstructure level. In order to check whether hysteretic models display spurious long–range dependence, a model possessing the hysteresis property is used for simulating hysteretic data to which statistical tests for short–memory against long–memory alternatives are applied.

We wish to thank all contributors to this book, all the referees for their careful and fast reports which made the selection of the chapters easier, the Springer Verlag editor, Dr Martina Bihn, for efficient and patient management, and Mrs Ruth Milewski for correcting the first series of manuscripts and the final version. Teyssière managed chapters 1 to 8 and 10, while Kirman managed chapters 9, 11, 12 and 13.

Gilles Teyssière thanks Paul Doukhan for his help to return to the academic system on serious grounds, and the SAMOS (Statistique Appliquée et MOdélisation Stochastique) University Paris 1. He wants to express his gratitude to Rosa Maria to whom he owes bunches of flowers as the completion of this volume spilt over into their spare time. Alan Kirman would like to thank the other people with whom he has investigated this sort of problem, and in particular, Hans Foellmer, Ulrich Horst, Richard Topol and Romain Ricciotti.

Paris,
June 2006

Gilles Teyssière
Alan Kirman

Contents

Part II Economic Models

Recent Advances in ARCH Modelling

Liudas Giraitis[1] ⋆, Remigijus Leipus[2], and Donatas Surgailis[3]

[1] University of York, Department of Mathematics and Economics. lg8@york.ac.uk
[2] Vilnius University, Institute of Mathematics and Informatics. Remigijus.Leipus@maf.vu.lt
[3] Vilnius Institute of Mathematics and Informatics. sdonatas@ktl.mii.lt

1 Introduction

Econometric modelling of financial data received a broad interest in the last 20 years and the literature on ARCH and related models is vast. Starting with the path breaking works by Engle (1982) and Bollerslev (1986), one of the most popular models became the Generalized AutoRegressive Conditionally Heteroscedastic (GARCH) process. The classical GARCH(p, q) model is given by equations

$$r_t = \sigma_t \varepsilon_t, \quad \sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 + \sum_{j=1}^{q} \alpha_j r_{t-j}^2, \tag{1}$$

where $\alpha_0 > 0, \alpha_j \geq 0, \beta_i \geq 0, \ p \geq 0, q \geq 0$ are model parameters and $\{\varepsilon_j, j \in \mathbf{Z}\}$ are independent identically distributed (i.i.d.) zero mean random variables. The variables r_t, σ_t, ε_t in (65) are usually interpreted as financial (log)returns (r_t), their volatilities or conditional standard deviations (σ_t), and so-called innovations or shocks (ε_t), respectively; in (65) the innovations are supposed to follow a certain fixed distribution (e.g., standard normal). Later, a number of modifications of (65) were proposed, which account for asymmetry, leverage effect, heavy tails and other "stylized facts". For statistical and econometric aspects of ARCH modelling, see the surveys of Bollerslev *et al.* (1992), Shephard (1996), Bera and Higgins (1993), Bollerslev *et al.* (1994); for specific features of modelling the financial data, including ARCH, see Pagan (1996), Rydberg (2000), Mikosch (2003). Berkes *et al.* (2002b) review some recent results. One should mention here, besides the classical reference to Taylor (1986), the related monographs by Gouriéroux (1997), Fan and Yao (2002), Tsay (2002). Let us note that the GARCH model for returns is also related to the Autoregressive Conditional Duration (ACD) model proposed by Engle and Russell (1998) for modelling of durations between events.

⋆ Supported by the ESRC grant R000239538.

Under some additional conditions, similarly as in the case of ARMA models, the GARCH model can be written as ARCH(∞) model (see (3) below), i.e., σ_t^2 can be represented as a moving average of the past squared returns $r_s^2, s < t$, with exponentially decaying coefficients (see Bollerslev, 1988) and absolutely summable exponentially decaying autocovariance function.

However, empirical studies of financial data show that sample autocorrelations of power series and volatilities (such as absolute values $|r_t|$ or squares r_t^2) remain non-zero for very large lags; see, e.g., Dacorogna *et al.* (1993), Ding *et al.* (1993), Baillie *et al.* (1996a), Ding and Granger (1996), Breidt *et al.* (1998), Mikosch and Stărică (2003), Andersen *et al.* (2001). These studies provide a clearcut evidence in favor of models with autocovariances decaying slowly with the lag as $k^{-\gamma}$, for some $0 < \gamma < 1$.

A number of such models (FIGARCH, LM-ARCH, FIEGARCH) were suggested in the ARCH literature. The long memory property was rigorously established for some of these models including the Gaussian subordinated stochastic volatility model (Robinson, 2001), with general form of nonlinearity, the FIEGARCH and related exponential volatility models (Harvey, 1998; Surgailis and Viano, 2002), the LARCH model (Giraitis *et al.*, 2000c), the stochastic volatility model of Robinson and Zaffaroni (1997, 1998). The long memory property (and even the existence of stationary regime) of some other models (FIGARCH, LM-ARCH) has not been theoretically established; see Giraitis *et al.* (2000a) Mikosch and Stărică (2000, 2003), Kazakevičius *et al.* (2004). Covariance long memory was also proved for some regime switching SV models (Liu, 2000; Leipus *et al.*, 2005). One should also mention that some authors (Mikosch and Stărică, 1999, 2004) argue that the observed long memory in sample autocorrelations can be explained by short memory GARCH models with structural breaks and/or slowly changing trends.

The present paper reviews some recent theoretical findings on ARCH type models. We focus mainly on covariance stationary models which display empirically observed properties known as "stylized facts". One of the major issues to determine is whether the corresponding model r_t^2 for squares has *long memory* or *short memory*, i.e. whether $\sum_{k=0}^{\infty} |\mathrm{Cov}(r_k^2, r_0^2)| = \infty$ or $\sum_{k=0}^{\infty} |\mathrm{Cov}(r_k^2, r_0^2)| < \infty$ holds. It is pointed out that for several ARCH-type models the behavior of $\mathrm{Cov}(r_k^2, r_0^2)$ alone is sufficient to derive the limit distribution of $\sum_{j=1}^{N}(r_j^2 - Er_j^2)$ and statistical inferences, without imposing any additional (e.g. mixing) assumptions on the dependence structure.

The review discusses ARCH(∞) processes and their modifications such as linear ARCH (LARCH), bilinear models, long memory EGARCH and stochastic volatility, regime switching SV models, random coefficient ARCH and aggregation. We give an overview of the theoretical results on the existence of a stationary solution, dependence structure, limit behavior of partial sums, leverage effect and long memory property of these models. Statistical estimation of ARCH parameters and testing for change-points are also discussed.

2 ARCH(∞)

A random sequence $\{r_t,\ t \in \mathbf{Z}\}$ is said to satisfy ARCH(∞) equations if there exists a sequence of i.i.d. zero mean random variables $\{\varepsilon_t,\ t \in \mathbf{Z}\}$ and a deterministic sequence $b_j \geq 0$, $j = 0, 1, \ldots$ such that for any t

$$r_t = \sigma_t \varepsilon_t,\ \sigma_t^2 = b_0 + \sum_{j=1}^{\infty} b_j r_{t-j}^2. \tag{2}$$

Clearly, if $E(\varepsilon_t | r_s, s < t) = 0$, $E(\varepsilon_t^2 | r_s, s < t) = 1$ then r_t has conditional mean zero and a random conditional variance σ_t^2, i.e.

$$E(r_t | r_s, s < t) = 0, \quad \mathrm{Var}(r_t | r_s, s < t) = \sigma_t^2.$$

The general framework leading to the model (2) was introduced by Robinson (1991) in the context of testing for strong serial correlation and has been subsequently studied by Kokoszka and Leipus (2000) in the change-point problem context. The class of ARCH(∞) models include the finite order ARCH and GARCH models of Engle (1982) and Bollerslev (1986). For instance, the GARCH(p, q) process $\{r_t, t \in \mathbf{Z}\}$ of (65) can be written as $r_t = \sigma_t \varepsilon_t$,

$$\sigma_t^2 = (1 - \beta(1))^{-1} \alpha_0 + (1 - \beta(L))^{-1} \alpha(L) r_t^2, \tag{3}$$

where $\beta(L) = \beta_1 L + ... + \beta_p L^p$ and L stands for the back-shift operator, $L^j X_t = X_{t-j}$. This leads to ARCH(∞) representation (2) for GARCH(p, q) model with $b_0 = (1 - \beta(1))^{-1} \alpha_0$ and with positive *exponentially* decaying weights $b_j, j \geq 1$ defined by the generating function $\alpha(z)/(1 - \beta(z)) = \sum_{i=1}^{\infty} b_i z^i$. It is interesting to note that the non-negativity of the regression coefficients α_j, β_j in (65) is not necessary for non-negativity of b_j in the corresponding ARCH(∞) representation, see Nelson and Cao (1992).

2.1 Existence of Second and Fourth Order Stationary Solutions

One of the first questions which usually arise in the study of recursion equations of the type (2) is to find conditions for the existence of a stationary solution. We first discuss conditions on the coefficients b_j and the random variables ε_t which guarantee the existence of a stationary solution to equations (2) with finite second or fourth moments.

Formally, recursion relations (2) give the following Volterra series expansion of r_t^2:

$$\begin{aligned} r_t^2 \equiv \varepsilon_t^2 \sigma_t^2 &= \varepsilon_t^2 b_0 \Big(1 + \sum_{k=1}^{\infty} \sum_{j_1, \ldots, j_k = 1}^{\infty} b_{j_1} \ldots b_{j_k} \varepsilon_{t-j_1}^2 \ldots \varepsilon_{t-j_1-\cdots-j_k}^2 \Big) \\ &= \varepsilon_t^2 b_0 \Big(1 + \sum_{k=1}^{\infty} \sum_{-\infty < s_k < \cdots < s_1 < t}^{\infty} b_{t-s_1} b_{s_1 - s_2} \ldots b_{s_{k-1} - s_k} \varepsilon_{s_1}^2 \ldots \varepsilon_{s_k}^2 \Big). \end{aligned} \tag{4}$$

By taking the expectation on both sides and using the independence of ε_t's, one obtains

$$Er_t^2 = (E\varepsilon_t^2)b_0\Big\{1+\sum_{k=1}^{\infty}\sum_{-\infty<s_k<\cdots<s_1<t}^{\infty} b_{t-s_1}b_{s_1-s_2}\dots b_{s_{k-1}-s_k}E\varepsilon_{s_1}^2\dots E\varepsilon_{s_k}^2\Big\}$$
$$= (E\varepsilon_t^2)b_0\Big\{1+\sum_{k=1}^{\infty}\Big(E\varepsilon_0^2\sum_{j=1}^{\infty}b_j\Big)^k\Big\} = \frac{b_0E\varepsilon_0^2}{1-E\varepsilon_0^2\sum_{j=1}^{\infty}b_j}.$$

Hence it easily follows that

$$E\varepsilon_0^2\sum_{j=1}^{\infty}b_j<1 \tag{5}$$

is sufficient for the existence of stationary solution (4) with $Er_t^2<\infty$. The uniqueness and the necessity of (5) for the existence of such a solution also follow easily, see Kokoszka and Leipus (2000), Giraitis *et al.* (2000a).

It is also easy to obtain a sufficient condition for the existence of a stationary solution with finite fourth moment. To that end, apply to (4) the norm (Minkowski) inequality: $(E\big(\sum_i\xi_i\big)^2)^{1/2}\le\sum_i(E\xi_i^2)^{1/2}$. Similarly as above, this yields

$$(Er_t^4)^{1/2}\le(E\varepsilon_t^4)^{1/2}$$
$$\times b_0\Big\{1+\sum_{k=1}^{\infty}\sum_{-\infty<s_k<\cdots<s_1<t}^{\infty} b_{t-s_1}b_{s_1-s_2}\dots b_{s_{k-1}-s_k}(E\varepsilon_{s_1}^4)^{1/2}\dots(E\varepsilon_{s_k}^4)^{1/2}\Big\}$$
$$=\frac{b_0(E\varepsilon_0^4)^{1/2}}{1-(E\varepsilon_0^4)^{1/2}\sum_{j=1}^{\infty}b_j}.$$

Hence if condition

$$(E\varepsilon_0^4)^{1/2}\sum_{j=1}^{\infty}b_j<1 \tag{6}$$

is satisfied, then r_t of (4) is a fourth order stationary solution to (2), see Giraitis *et al.* (2000a). A similar norm inequality works in the case of $E(r_t^2)^p$ and arbitrary $p\ge1$, yielding a sufficient condition $(E|\varepsilon_0|^{2p})^{1/p}\ \sum_{j=1}^{\infty}b_j<1$.

Condition (6) is not necessary for the existence of fourth order stationary solution. For example, in the case of GARCH(1,1) $r_t=\varepsilon_t\sigma_t$, $\sigma_t^2=\alpha_0+\alpha r_{t-1}^2+\beta\sigma_{t-1}^2$, (6) translates to $\alpha\lambda_2^{1/2}+\beta<1$, $\lambda_i=E\varepsilon^{2i}, i=1,2$, while a fourth order stationary solution is known to exist under the weaker conditions

$$\alpha\lambda_1+\beta<1,\qquad \alpha^2\lambda_2+\beta^2<1, \tag{7}$$

see Karanasos (1999), He and Teräsvirta (1999). To obtain a sufficient and necessary condition in the general case, one needs to study *orthogonal* Volterra representation of r_t^2.

The orthogonal Volterra representation (9) of r_t^2 is obtained by centering the innovations in the (nonorthogonal) representation (4), i.e. by replacing the ε_j^2's by $\kappa\zeta_j + \lambda_1 = \varepsilon_j^2$, where the standardized $\zeta_j = (\varepsilon_j^2 - E\varepsilon_j^2)/\kappa$, $\kappa^2 = \mathrm{Var}(\varepsilon_0^2)$ have zero mean and unit variance.

The resulting expression appears rather complicated, but nevertheless it can be identified and studied (Giraitis and Surgailis, 2002). In order to describe it, denote g_j the coefficients of the generating function

$$\sum_{j=0}^{\infty} g_j z^j = \left(1 - \lambda_1 \sum_{i=1}^{\infty} b_i z^i\right)^{-1}.$$

More explicitly,

$$g_j = \sum_{k=1}^{j} \lambda_1^k \sum_{0<i_1<\ldots<i_{k-1}<j} b_{i_1} b_{i_2-i_1} \ldots b_{i_{k-2}-i_{k-1}} b_{j-i_{k-1}} \quad (j \geq 1), \quad g_0 = 1. \tag{8}$$

Also introduce $h_j = (\kappa/\lambda_1) g_j,\ j \geq 1$,

$$B = \sum_{j=1}^{\infty} b_j, \qquad H^2 = \sum_{j=1}^{\infty} h_j^2.$$

Then

$$r_t^2 = \mu + (\kappa/\lambda_1)\mu \sum_{k=1}^{\infty} \sum_{s_k<\cdots<s_2<s_1\leq t} g_{t-s_1} h_{s_1-s_2} \ldots h_{s_{k-1}-s_k} \zeta_{s_1} \ldots \zeta_{s_k}, \tag{9}$$

where $\mu = Er_t^2 = \lambda_1 b_0/(1-\lambda_1 B)$. The series (9) converges in mean square if and only if

$$\lambda_1 B < 1, \qquad H < 1 \tag{10}$$

hold, and define a stationary solution of (2).

In fact, conditions (10) are sufficient and necessary for the existence of fourth order stationary solution of (2) (Giraitis and Surgailis, 2002). By orthogonality, it easily follows that

$$\begin{aligned}
\mathrm{Cov}(r_t^2, r_0^2) &= (\kappa/\lambda_1)^2 \mu^2 \sum_{k=1}^{\infty} \sum_{s_k<\cdots<s_1\leq 0} g_{-s_1} g_{t-s_1} h_{s_1-s_2}^2 \ldots h_{s_{k-1}-s_k}^2 \\
&= (\kappa/\lambda_1)^2 \mu^2 \sum_{s\leq 0} g_s g_{t-s} \sum_{k=1}^{\infty} H^{2(k-1)} \\
&= \frac{(\kappa/\lambda_1)^2 \mu^2}{1-H^2} \sum_{s=0}^{\infty} g_s g_{s+t}.
\end{aligned} \tag{11}$$

For the GARCH(1,1) model

$$r_t = \varepsilon_t \sigma_t, \quad \sigma_t^2 = \alpha_0 + \alpha r_{t-1}^2 + \beta \sigma_{t-1}^2$$

the above formulas are more explicit. The model itself can be rewritten in the ARCH(∞) form

$$\sigma_t^2 = \alpha_0 (1-\beta)^{-1} + \alpha \sum_{j=1}^{\infty} \beta^{j-1} r_{t-j}^2.$$

In this case, $g_j = \alpha\lambda_1(\lambda_1\alpha + \beta)^{j-1}$, conditions (7) and (10) coincide, and (9) becomes

$$r_t^2 = \mu + \mu\kappa\lambda_1^{-1}\zeta_t\Big(1 + \sum_{k=1}^{\infty}(\alpha\kappa/\gamma)^k \sum_{s_k<\cdots<s_1<t} \gamma^{t-s_k}\zeta_{s_1}\ldots\zeta_{s_k}\Big)$$
$$+ \mu\sum_{k=1}^{\infty}(\alpha\kappa/\gamma)^k \sum_{s_k<\cdots<s_1<t} \gamma^{t-s_k}\zeta_{s_1}\ldots\zeta_{s_k}, \tag{12}$$

where $\gamma = \lambda_1\alpha + \beta$, $\mu = \alpha_0\lambda_1/(1-\gamma)$. From (12) or (11) one can explicitly find the variance and covariance function of the GARCH(1,1) model in terms of the coefficients α_0, α, β and the moments λ_1, λ_2:

$$\operatorname{Var} r_0^2 = \frac{\alpha_0^2\kappa^2(1-\gamma^2+\gamma\alpha\lambda_1)}{(1-\gamma)^2(1-\gamma^2-\alpha^2\kappa^2)},$$
$$\operatorname{Cov}(r_k^2, r_0^2) = \frac{\alpha_0^2\alpha\lambda_1\kappa^2(1-\gamma^2+\gamma\alpha\lambda_1)}{(1-\gamma)^2(1-\gamma^2-\alpha^2\kappa^2)}\,\gamma^{k-1}, \quad k \geq 1,$$

which were also obtained in Teräsvirta (1996). We also note an alternative approach in Kazakevičius *et al.* (2004) to the problem of the existence of fourth order stationary solution of ARCH(∞), which leads to equivalent necessary and sufficient conditions as (10).

For necessary and sufficient conditions of the existence of high order moments for the family of GARCH processes see Ling and McAleer (2002a, 2002b). Ling and McAleer (2003a) studied theoretical properties of the multivariate ARMA-GARCH model.

2.2 Dependence Structure, Association and Limit Theorems

The equation (11) for the covariance of ARCH(∞) squares r_t^2 allows to directly study its summability and decay properties. From (8) and the summability of b_j's it follows the summability of g_j's which in turn implies by (11) the summability of the autocovariances of r_t^2:

$$\sum_{k=-\infty}^{\infty} \operatorname{Cov}(r_k^2, r_0^2) < \infty. \tag{13}$$

(Note that $\operatorname{Cov}(r_k^2, r_0^2) \geq 0$ for all k, which follows from (11) and also from the associativity property of r_t^2, see below.) Equation (13) indicates that the *squares r_t^2 of a fourth order stationary solution of ARCH(∞) have short memory.* The above mentioned papers Giraitis *et al.* (2000a), Giraitis and Surgailis (2002) also prove that a hyperbolic decay $b_j \sim Cj^{-\gamma}$ with $\gamma > 1$ implies

$$\mathrm{Cov}(r_k^2, r_0^2) \asymp k^{-\gamma}.$$

(Here and below, $x_k \sim y_k$ means $x_k/y_k \to 1$ while $x_k \asymp y_k$ means that there are positive constants C_1 and C_2 such that $C_1 y_k < x_k < C_2 y_k$ for all k.) Thus, even though condition (6) implies absolute summability of the covariances, it allows for a very slow rate of decay of the autocorrelation function when $\gamma > 1$ is close to 1. The last property may be characterized as *moderate* memory. Near epoch dependence and moderate memory property of the so-called HYGARCH model were studied by Davidson (2004).

The above discussion basically concerns second-order properties of r_t^2 only. Some further insight about these properties can be obtained from the *moving average representation*

$$r_t^2 = Er_t^2 + \sum_{j=0}^{\infty} g_j \nu_{t-j}, \qquad (14)$$

where g_j (8) and $\nu_t \equiv \sigma_t^2(\varepsilon_t^2 - E\varepsilon_t^2)$ are *martingale differences.* The above representation is a direct consequence of (9), from which the ν_t's can be also expressed as a Volterra series in the standardized variables $\zeta_s, s < t$. Of course, (14) yields the same covariance formula as (11). On the other hand, the ν_t's are *not* independent, meaning that "higher order" dependence and distributional properties of (14) may be very different from the usual moving average in i.i.d. random variables.

ARCH(∞) sequences have important property of associativity. A random sequence $\{X_t\}$ is said to be *associated* (or *positively correlated*) if the inequality

$$\mathrm{Cov}(f(X_{t_1}, \ldots, X_{t_n}), g(X_{t_1}, \ldots, X_{t_n})) \geq 0,$$

holds for any coordinate nondecreasing functions $f, g : \mathbf{R}^n \to \mathbf{R}$ and any $t_1, \ldots, t_n, n = 1, 2, \ldots$. In particular, the covariance function (if it exists) of associated sequence is nonnegative: $\mathrm{Cov}(X_s, X_t) \geq 0$ for any s, t. Associated sequences are widely encountered in applications, see e.g. Barlow and Proschan (1981), Newman (1984), Cox and Grimmett (1984). Association is a very strong property, under which uncorrelatedness implies independence similarly as in Gaussian case. A number of limit theorems have been proved for associated sequences under covariance restrictions only. One of the most celebrated results, due to Newman and Wright (1981), says that if $\{X_t, t \in \mathbf{Z}\}$ is strictly stationary and associated and $\sigma^2 = \sum_{t \in \mathbf{Z}} \mathrm{Cov}(X_0, X_t) < \infty$ then the partial sums' process

$$\left\{ N^{-1/2} \sum_{t=1}^{[N\tau]} (X_t - EX_t), \tau \in [0,1] \right\} \to_{D[0,1]} \{\sigma W(\tau), \tau \in [0,1]\}, \qquad (15)$$

in the Skorokhod space $D[0,1]$, where $\{W(\tau)\}$ is a standard Brownian motion.

It is well known that independent random variables are associated, and that this property is preserved by coordinate-nondecreasing (nonlinear) transformations. In particular, the ARCH(∞) process of (4) is a coordinate-nondecreasing transformation of the i.i.d. sequence $\{\varepsilon_t^2\}$. It is clear from non-negativity $b_j \geq 0,\ j \geq 0$ that r_t^2 can only increase if any of $\varepsilon_s^2, s \leq t$ on the r.h.s. of (4) is replaced by some larger quantity. Therefore the *ARCH(∞) process* (4) *is associated.*

An immediate consequence of (15), (13) and the association property of is the functional central limit theorem for squares r_t^2 of ARCH(∞):

$$\left\{ N^{-1/2} \sum_{t=1}^{[N\tau]} (r_t^2 - Er_t^2),\ \tau \in [0,1] \right\} \to_{D[0,1]} \ \{\sigma W(\tau), \tau \in [0,1]\}, \tag{16}$$

where σ^2 equals to the sum in (13). This result is quite surprising given a rather complicated nonlinear structure of ARCH(∞), since it holds for any stationary solution r_t such that $Er_t^4 < \infty$. Giraitis *et al.* (2000a) obtained a similar result by using finite memory approximation to ARCH(∞).

It seems that the implications of association property to the study of ARCH models have been not yet fully explored. This remark applies e.g. to the covariance structure and dependence properties of general nonlinear transformations of ARCH(∞), Rosenthal inequalities, rate of convergence, empirical processes and many other questions. See the dissertation of Louichi (1998) for references.

2.3 Stationary Solution of ARCH(∞) Without Moment Assumptions

A rather unusual feature of ARCH equations is the fact that they may admit a stationary solution which does not have any moments, even if the i.i.d. "shocks" ε_t's are $N(0,1)$. In such case, the Volterra series (4) converge in probability but not in any moment sense, and the properties of the infinite series are much more difficult to study. Nelson (1990) showed, using the theory of products of random matrices, that a necessary and sufficient condition for the existence of a strictly stationary GARCH(1,1) process is

$$E \log(\alpha\varepsilon_0^2 + \beta) < 0. \tag{17}$$

This condition is of course much weaker than any of conditions (5), (6), (7) given above (which imply in particular the existence of finite moment $Er_t^2 < \infty$ or $Er_t^4 < \infty$). Nelson's result was extended to the GARCH(p,q) case by Bougerol and Picard (1992), who showed that, under condition $E\varepsilon_0^2 = 1$, a stationary solution to (65) exists if and only if the top Lyapunov exponent γ is strictly negative. Moreover, in such case there exists only one stationary GARCH(p,q) process. The top Lyapunov exponent is defined by

$$\gamma = \lim_{n \to \infty} n^{-1} \log \|A_1 \cdots A_n\|, \tag{18}$$

where $\{A_k\}$ are i.i.d. $(p+q-1) \times (p+q-1)$ random matrices (depending only on ε_k^2) such that the $(p+q-1)$-valued process $X_t = (\sigma_{t+1}^2, \ldots, \sigma_{t-p+2}^2, r_t^2, \ldots, r_{t-q+2}^2)'$ satisfies the random coefficient matrix AR(1) equation

$$X_t = A_t X_{t-1} + B,$$

with $B = (\alpha_0, 0, \ldots, 0)'$; see Bougerol and Picard (1992).

Further progress in this direction was made by Kazakevičius and Leipus (2002), who discussed the general case of ARCH(∞). They observed that the volatility can be written as

$$\sigma_t^2 = b_0 \left(1 + \sum_{n=1}^{\infty} \sigma_{t,n}^2\right), \tag{19}$$

the convergence of the series being equivalent to the existence of a stationary solution $r_t = \varepsilon_t \sigma_t$, where

$$\sigma_{t,n}^2 = \sum_{k=1}^{n} \sum_{\substack{i_1, \ldots, i_k \geq 1 \\ i_1 + \cdots + i_k = n}} b_{i_1} b_{i_2} \cdots b_{i_k} \varepsilon_{t-i_1}^2 \varepsilon_{t-i_1-i_2}^2 \cdots \varepsilon_{t-i_1-\cdots-i_k}^2 \quad (n \geq 1),$$

$\sigma_{t,n}^2 = 0$ $(n \leq 0)$, satisfy the recurrent equation

$$\sigma_{t,n}^2 = \varepsilon_{t-n}^2 \sum_{i=1}^{n} b_i \sigma_{t,n-i}^2, \quad n \geq 1. \tag{20}$$

Equation (20) is a "stochastic" version of the corresponding equation satisfied by $g_n = E\sigma_{t,n}^2$ of (8) in the case $\lambda_1 = E\varepsilon_0^2 < \infty$. For a fixed t (say, $t = 0$), equation (20) can be written in the matrix form:

$$\Sigma_n = B_n \Sigma_{n-1},$$

where $\Sigma_n = (\sigma_{0,n}^2, \sigma_{0,n-1}^2, \ldots, \sigma_{0,1}^2, 0, \ldots)'$ and where $\{B_n\}$ are random i.i.d. (infinite) matrices:

$$B_n = \begin{pmatrix} b_1 \varepsilon_{-n}^2 & b_2 \varepsilon_{-n}^2 & b_3 \varepsilon_{-n}^2 & \cdots \\ 1 & 0 & 0 & \cdots \\ 0 & 1 & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix}. \tag{21}$$

Kazakevičius and Leipus (2002) showed that in the GARCH(p, q) case, the top Lyapunov exponent satisfies

$$\gamma = -\log R, \tag{22}$$