CENTRE FOR ECONOMETRIC ANALYSIS
CEA@Cass

http://www.cass.city.ac.uk/cea/index.html

Cass Business School
Faculty of Finance
106 Bunhill Row
London EC1Y 8TZ

Modelling Structural Breaks, Long Memory and Stock Market Volatility: An Overview

Anindya Banerjee and Giovanni Urga

CEA@Cass Working Paper Series

WP–CEA–07-2004

Modelling structural breaks, long memory and stock market volatility: an overview

Anindya Banerjee
Department of Economics, European University Institute,
Villa San Paolo, Via della Piazzuola 43, 50133 Firenze, Italy

Giovanni Urga*†
Faculty of Finance, Cass Business School, London,
106, Bunhill Row, London, EC1Y 8TZ, U.K.

28 August 2004

Abstract

The main aim of this volume is to present key recent developments in the fields of modelling structural breaks, and the analysis of long memory and stock market volatility.

Keywords: Structural Breaks, Long Memory, Volatility.
JEL Classification: C22.

*Corresponding author. Tel.: +44-20-70408698; fax. +44-20-70408881.
†E-mail addresses: anindya.banerjee@iue.it. (A. Banerjee), g.urga@city.ac.uk (G.Urga).

1 Introduction

This special Annals issue of the Journal of Econometrics contains a collection of papers from the conference on "Long Memory, Structural Breaks and Stock Market Volatility", organized by us in London at the Cass Business School from the 5th to the 7th of December 2002.

Modelling strong persistence in time series has constituted a major research agenda in the econometrics literature for a number of years. One approach has been to look at the long-memory properties of time series and to study mechanisms which generate series having these properties, such as the various forms of fractionally integrated, autoregressive moving average and non-linear models. In the macroeconometrics literature, attention has focused more on looking at the properties of stochastic processes with unit roots. Unit root processes can be viewed as a specific form of long memory. A significant strand of the debate has also considered the properties of tests for unit root, cointegration or long memory in the presence of structural breaks. It has been shown that persistent tests are severely compromised, in terms of their size and power properties, in series which display breaks, particularly in their deterministic components, because these processes give the impression of persistence. These issues all have wider reach and the interface between the literatures on unit roots, structural breaks and long memory is often important.

Our editorial introduction is in two parts and consists of discussing the developments in the literatures on structural breaks and long memory, with specific applications. Our goal is to obtain an overall sense of the main frameworks under which research in these areas has so far been conducted. The papers in this volume, belonging to each area, are then placed within these frameworks and their contributions are discussed. The invited keynote address by Nobel Laureate Professor Clive Granger proposes some key areas for further developments in the literature.

Although based on considering a vast number of papers, our approach in writing this introduction is necessarily selective. However, during the course of preparation of this volume, a key resource for econometricians has been the compilation of an archive contained in a CD-Rom (available freely from us upon request) of the majority of recently published papers in these areas. Details of papers not included in this introduction may thus be consulted with ease.

2 Structural breaks

A rich literature on the econometrics of structural breaks has developed in recent years, and an earlier annals volume of this journal published in 1996 entitled *Recent Developments in the Econometrics of Structural Change* (edited by Jean-Marie Dufour and Eric Ghysels) surveyed some of the significant developments in the field of structural change at that time.

Any description of the 'problem' of dealing with structural breaks (and, more generally, structural change) in estimation and inference may be organised in several different ways, in order to encompass the cases characterized by considering linear or non-linear models, stationary or non-stationary regressors, known or unknown point of break, multiple breaks or single break, estimation in single-equation or systems, or any of these in combination.

In the following sections we describe three papers which help to illustrate some of the issues involved. As background, we start with a brief description of the debate on whether macroeconomic time series are characterized by breaks in their deterministic components, which developed in response to Nelson and Plosser's (1982) powerful critique which argued in favour of unit roots in many of these series.

2.1 Structural Breaks - a background

In 1960s and 1970s business cycles were commonly thought of as departures from a secular trend. Consequently, the time series of macroeconomic variables were usually decomposed into a trend component and a cyclical component. The trend was believed to be deterministic, often linear, while departures were assumed to be stationary, therefore transitory.

Nelson and Plosser (1982) were the first to point out that the secular component need not be modelled by a deterministic trend and that a possible stochastic nature of the trend should be considered. They examined several long US macroeconomic time series and were unable to reject the null of a unit root for twelve out of fourteen series in their data set.

Their finding had a profound impact on the way economic series have been viewed and treated from then on. First, the results confirmed previous analyses that suggested that some specific variables followed a random walk, for example stock prices (Samuelson, 1973) or consumption (Hall, 1979). Secondly, it spurred intensive theoretical investigation of unit root processes and led to the definition and examination of cointegration among several variables (Engle and Granger, 1987, Stock and Watson, 1988, Phillips and Durlauf, 1986). Perhaps the most important implication for the theory of business cycles was that if the series were indeed integrated, random shocks would have a permanent effect on the economy.

In two seminal works, responding to the unit root agenda, Rappoport and Reichlin (1989) and Perron (1989) argued that the majority of shocks to the key economic variables of any economywould be transitory and that only few events would have any permanent effect. They represented such shocks as breaks in the underlying deterministic trends. They demonstrated that if structural breaks were present in the data generating process but not allowed for in the specification of an econometric model, the analysis would be biased towards erroneous non-rejection of the unit root hypothesis. Upon re-examination of the Nelson-Plosser data set, under a specification of the model which allowed for a single break in the deterministic components of the series, Perron was able to reject the unit root hypothesis in twelve out of fourteen series. Similarly, Rappoport and Reichlin were not able to confirm Nelson and Plosser's results, when allowing for break in mean or trend.

The work of Rappoport, Reichlin and Perron was subsequently criticised in several papers published in the Journal of Business and Economic Statistics in 1992. Zivot and Andrews (1992) *inter alia* pointed out that the specification and choice of breakpoint in Perron (1989) was influenced by a prior examination of the data. These papers therefore argued in favour of the need to view break points as endogenous and to develop procedures which took this endogeneity into account.

The subsequent development of the literature on testing for unit roots thus took place in two directions - the development of procedures which (unlike Perron, 1989) estimated the date of break or structural change and secondly, procedures which allowed for the possibility of multiple breaks or changes during the lifetime of the sample. We explore these two directions in turn in the following two sub-sections.

2.2 Estimating the date of break (Zivot and Andrews, 1992)

Consider the so-called "changing growth" model of Perron (1989) where the null hypothesis given by the following Model B (in Perron's terminology) is

$$y_t = \mu_1 + y_{t-1} + (\mu_2 - \mu_1)DU_t + e_t,$$

where $DU_t = 1$ if $t > T_B$, 0 otherwise and e_t is an ARMA (p, q) process. Under the trend-stationary alternative hypothesis,

$$y_t = \mu_1 + \beta_1 t + (\beta_2 - \beta_1)DT_t^* + e_t,$$

where $DT_t^* = t - T_B$ if $t > T_B$ and is zero otherwise. Therefore, under the null, the series has a unit root with drift and an exogenous structural break occurs at $1 < T_B < T$. Under the alternative hypothesis, the series is stationary around a deterministic trend with a break in the trend function that occurs at time exogenously given T_B.

Perron's estimated model, corresponding to this specification of changing growth, is given by

$$y_t = \widehat{\mu^B} + \widehat{\beta}_B t + \widehat{\gamma}_B DT_t^* + \widehat{\alpha}^B y_{t-1} + \sum_{j=1}^{k} \widehat{c}_j^B \Delta y_{t-j} + \widehat{e}_t,$$

and to test formally for the presence of a unit root requires the construction of t-statistics given by

$$t_{\widehat{\alpha}^B}(\lambda)$$

where λ is the known fraction of the sample (T_B/T) at which the break in the trend function occurs. The dependence of the $t-$statistic on the break fraction arises through through the construction of the variable DT_t^* and also through the dependence of the critical values used to test the unit root hypothesis (H_0 :

$\alpha_B = 1$) on λ - i.e. the critical values are a function of where the break is assumed to occur. Perron's method thus corresponds to testing for a unit root with (possible) structural change of known form (abruptly or continuously, corresponding to additive and innovation outlier models respectively) in the deterministic growth rate of the variable.

Allowing the break date to be unknown leads to the consideration of statistics discussed by Banerjee, Lumsdaine and Stock (1992) and Zivot and Andrews (1992) of the form

$$\inf_{\lambda \in \Lambda} t_{\widehat{\alpha}^B}(\lambda)$$

where Λ is a pre-defined closed subset of $(0,1)$ but λ is assumed unknown. Thus $t-$statistics are computed, with the break date T_B allowed to vary across the length of the sample (located at $[T\lambda]$ for $\lambda \epsilon \Lambda$) with no need for trimming as proved by Perron (1997). The subsequent $\inf_{\lambda \in \Lambda} t_{\widehat{\alpha}^B}(\lambda)$ is then compared against the relevant critical values computed by Zivot and Andrews. This form of test statistic gives the greatest weight to the trend-break hypothesis under the alternative.

Other such summary or order statistics can be considered to allow for unknown date of break, such as those postulated by Hansen (1992), which have been put to popular use in many empirical applications.

2.3 Making it more general - Multiple breaks

It is natural to think of the description of the problem above as being too restrictive, even allowing for the possibility that the date of break is unknown. There is no reason to consider the problem solely within the context of testing for a unit root with structural breaks. More particularly, it is also natural to allow for multiple changes. Once the issues arising in estimating or detecting one break were better understood, the attention naturally turned to the problem of multiple breaks. Just as a model of a process with a change of parameters would be misspecified if the presence of break was ignored, allowing for one break when in fact multiple breaks are present could lead to false conclusions.

We introduce the arguments in this section in two stages, beginning with a description of the multiple-change model in single equations as developed by Bai and Perron (1998) (henceforth BP), not allowing for stochastically trending regressors. We then move on to considering estimation in systems, where the regressors themselves may be subject to structural change, and allowing for stochastically trending regressors. The latter development is due to Hansen (2000) and contains almost all the empirically relevant cases that are of interest - multiple changes, stochastically trending regressors and structural change in systems.

For single equation estimators, following Andrews (1993), we may think of a general framework of analysis by considering a class of parametric models indexed by parameters (β, δ_t), $t = 1, 2, ...$ The null hypothesis is of stability:

$$H_0 : \delta_t = \delta_0, \text{for all } t \geqq 1, \text{for some } \delta_0 \in D \subset R^p.$$

The division of the parameter space into β and δ_t, as described in more detail below, reflects the distinction between 'pure' and 'partial' structural change models. In the former, β does not appear (i.e. all parameters in the model are subject to structural change) while in the latter case β appears and is held constant both under the null and alternative hypotheses.

The alternative hypotheses considered can have many forms. For a one-time change, occurring at a fraction λ (equivalently at time $T\lambda$, where T is the length of the sample), the alternative takes the form

$$\begin{aligned} H_{1T}(\pi) &: \\ \delta_t &= \delta_{1\pi}, \qquad t = 1, ...T\pi \\ \delta_t &= \delta_{2\pi}, \qquad t = T\pi + 1, ...T \end{aligned}$$

for some constants $\delta_{1\pi}, \delta_{2\pi} \in D \subset R^p$.

In the case of multiple changes, a more detailed partition of the sample into m regimes (with $\delta_t = \delta_{j\pi}$, for $j = 1, ..m$) is considered. The dates of the changes may be assumed to be known or unknown, and the form of the changes considered may be restricted to occur, for example, in only the deterministic components of the model. The changes themselves may occur as discrete breaks or more gradually and the model specifications are allowed to be quite diverse.

The BP procedure for testing for multiple structural breaks at unknown dates can be nicely structured within this framework. The model considered is a multiple linear regression model with m breaks ($m + 1$ regimes):

$$y_t = x_t'\beta + z_t'\delta_j + u_t, \qquad t = T_{j-1} + 1, \ldots, T_j \tag{1}$$

for $j = 1, \ldots, m + 1$, where by convention $T_0 = 0$ and $T_{m+1} = T$. Here, y_t is the observed independent variable, $x_t : (p \times 1)$ and $z_t : (q \times 1)$ are vectors of covariates, and β and $\delta_j : (j = 1, \ldots, m + 1)$ are the corresponding vectors of coefficients.

The break points $(T_1, \ldots, T_m)$ are treated as unknown. When β is not subject to shifts and is effectively estimated using the entire sample, the model is a *partial change model* in the terminology above. If $p = 0$, the model is called a *pure structural change model* where all the coefficients are subject to change.

The task is first to estimate consistently $(\beta^0, \delta_1^0, ...\delta_m^0, T_1^0, \ldots, T_m^0)$,where the '0' index denotes the true or null values of the parameters, and then to test for the presence of structural change.

The method of estimation is based on least squares, where for each $m-$partition of the sample $(T_1, \ldots, T_m)$, estimates of the parameters are obtained by minimising the sum of squared residuals (SSR) from (1) above. The estimated break points $(\widehat{T_1}, \ldots, \widehat{T_m})$ are given as the outcome of the algorithm

$$(\widehat{T_1}, \ldots, \widehat{T_m}) = \arg \min_{T_1,\ldots,T_m} S_T(T_1, \ldots, T_m)$$

where $S_T(T_1, \ldots, T_m)$ is the SSR from (1) for a given $m-$partition and the minimisation is taken over all partitions $(T_1, \ldots, T_m)$ such that $T_i - T_{i-1} \geqq q$.

During the course of a lengthy and rich paper, BP detail several interesting estimating and testing issues. Their first test is a *sup F* type test of no structural break against the alternative of a known and fixed number of $m = k$ structural breaks. The partition is defined as $T_i = [T\lambda_i]$, thereby determining the fractions in the sample at which the breaks occur as $(\lambda_1, ...\lambda_k)$.

Defining a set for an arbitrary small positive number ε :

$$\Lambda_\varepsilon = \{(\lambda_1, ...\lambda_k); \mid \lambda_{i+1} - \lambda_i \mid \geqq \varepsilon, \lambda_1 \geqq \varepsilon, \lambda_k \leq 1 - \varepsilon\},$$

to allow for the breaks to be asymptotically distinct and not to occur too near the end-points of the sample, the *sup F* test is then defined as

$$\sup F_T(k; q) = \sup_{(\lambda_1, ...\lambda_k) \in \Lambda_\varepsilon} F_T(\lambda_1, \ldots, \lambda_k; q).$$

This test is a generalisation of the *sup F* test considered by Andrews (1993) for $k = 1$ The *sup F* test will have a limit distribution that depends on heterogeneity and serial correlation if no account is made to correct for them. But if such features are viewed as a possibility then a corrected version, as defined in BP and in Bai and Perron (2003a, b) should be used, which has a limit distribution free of nuisance parameters. BP present critical values to cover up to 9 breaks and up to 10 regressors z_t whose coefficients are subject to change.

Several generalisations are also discussed. The first generalisation relates to relaxing the assumption that the number of breaks is known by allowing for an unknown number of breaks under the alternative given an upper bound M. The *Double Maximum* test is then defined as:

$$D \max F_T(M, q, a_1, \ldots, a_M) = \max_{1 \leq m \leq M} a_m \sup_{(\lambda_1, \ldots, \lambda_m) \in \Lambda_\varepsilon} F_T(\lambda_1, \ldots, \lambda_m; q) \quad (2)$$

defined for some weights $\{a_1, \ldots, a_M\}$ to reflect priors on the likelihood of various numbers of breaks.

By choosing all the weights equal to unity BP define the $UDmax$ as:

$$UD \max F_T(M, q) = \max_{1 \leq m \leq M} \sup_{(\lambda_1, \ldots, \lambda_m) \in \Lambda_\varepsilon} F_T(\lambda_1, \ldots, \lambda_m; q) \quad (3)$$

Since for any fixed q the critical values of the individual test $\sup_{(\lambda_1, \ldots, \lambda_m) \in \Lambda_\varepsilon} \mathrm{F}_T(\lambda_1, \ldots, \lambda_m; \mathrm{q})$ decreases as m increases, this implies that the marginal p-values decrease with m and may lead to a low power if the number of breaks is large. To alleviate this problem, BP also propose the $WD \max$ test that considers a set of weights such that the marginal p-values are equal across values of m. This implies weights that depends on q and the significance level of the test.

Let $c(q, \alpha, m)$ be the asymptotic critical value of the test $\sup_{(\lambda_1, \ldots, \lambda_m) \in \Lambda_\varepsilon} F_T(\lambda_1, \ldots, \lambda_m; q)$ for a significance level of α, and $a_1 = 1$ and $a_m = c(q, \alpha, 1)/c(q, \alpha, m)$ for $m > 1$, the $WD \max$ is defined as:

$$WD\max F_T(M,q) = \max_{1\le \mathrm{m}\le \mathrm{M}} \frac{\mathrm{c(q,}\alpha\mathrm{,1)}}{\mathrm{c(q,}\alpha\mathrm{,m)}} \times \sup_{(\lambda_1,\ldots,\lambda_m)\in\Lambda_\varepsilon} F_T(\lambda_1,\ldots,\lambda_m;q) \tag{4}$$

BP present critical values for $M = 5$ for both the UD and WD tests. These have recently been extended in Bai and Perron (2003b) to cover a wider range of values of ε.

The second set of generalisations considered by them is to test for the null hypothesis of l breaks against the alternative that an additional break exists, where the total number of breaks may be unknown. Their strategy consists of testing each so-called $l-$partition (obtained using the estimated partition $(\widehat{T}_1,\ldots,\widehat{T}_l)$) for the presence of an additional break. A fuller development of this sequential procedure is given by Bai (1999) as described below, which does not require conditioning upon previously obtained break dates to proceed and allows a direct comparison of the sum of squares obtained with l breaks with that obtained with $l+1$ breaks.

2.3.1 Bai's Likelihood Ratio Test (1999)

Bai (1999) proposed a likelihood ratio test for detecting multiple structural breaks with the null of l breaks against the alternative $l+1$ break points. This method allows optimal estimation under both null and alternative hypotheses simultaneously, in contrast with Bai-Perron test that requires conditioning on the previously obtained break dates. Using the same notation as above for the SSR for a given $m-$partition , the test statistic is based on the difference between the optimal SSR with l breaks and the optimal SSR with $l+1$ breaks and defined as

$$\sup LR_T(l+1|l) = \frac{S_T\left(\widehat{T}_1,\ldots,\widehat{T}_l\right) - S_T\left(\widehat{T}_1^*,\ldots,\widehat{T}_{l+1}^*\right)}{S_T\left(\widehat{T}_1^*,\ldots,\widehat{T}_{l+1}^*\right)/T} \tag{5}$$

where $\left(\widehat{T}_1,\ldots,\widehat{T}_l\right)$ is the estimator of the break point $\left(T_1^0,\ldots,T_l^0\right)$ under the null and$\left(\widehat{T}_1^*,\ldots,\widehat{T}_{l+1}^*\right)$ is the point at which the SSR is minimized when an additional break is allowed.

Critical values are obtained analytically since the limiting distribution has a known density function given by

$$\lim_{T\to\infty} P\left(\sup LR_T(l+1|l) > c\right) = 1 - \prod_{i=1}^{l+1}\left(1 - G_i(c)\right) \tag{6}$$

where

$$G_i(c) = \frac{c^{q/2}\exp(-c/2)}{2^{q/2-1}\Gamma(q/2)}\left[\left(1-\frac{q}{c}\right)\log\frac{1-\eta_i}{\eta_i} + \frac{2}{c} + o(c^{-2})\right].$$

q is the dimension of the z_t vector as above and $\eta_i = \frac{\varepsilon}{\lambda_i^o - \lambda_{i-1}^0}$. Since in practice λ_i^o or λ_{i-1}^0 may not be known, these may be replaced by their estimated sample equivalents given by $\widehat{\lambda_i^o}$ and $\widehat{\lambda_{i-1}^0}$ in the construction of the statistic. For given significance level of test α, the corresponding value of c can be computed. The test is shown to be consistent for the number and location of the break points and can be implemented by the use of numerical optimisation algorithms due to Bai and Perron (1998, 2003a, 2003b).

2.4 Hansen (2000)

We turn finally in this section to considering a formulation of the exercise that takes many of the above features into account, to allow for trending and non-trending (stochastic or deterministic) regressors (which would help to address the unit root versus structural break debate), multiple versus single breaks and known versus unknown break points. A significant point of departure not referred to above but considered here, concerns allowing for changes in the explanatory variables in (1). Hansen (2000) points out that it is usually unrealistic to exclude structural change in the marginal distributions. He proposes a model consisting of a conditional and several marginal processes. In so doing, he is able to distinguish conceptually between the question of changes of parameters in regression and the question of structural stability of the regressors.

The linear regression model considered by Hansen is

$$y_t = x'_{nt}\beta_{nt} + e_{nt}, \quad t = 1, \ldots, T. \tag{7}$$

with $E(e_{nt}^2) = \sigma^2 < \infty$. If there is a structural change, β takes the form of

$$\beta_{nt} = \begin{cases} \beta, & , t < T_0 \\ \beta + \theta_n, t \geq T_0 \end{cases} \tag{8}$$

where $T_0 \in \begin{bmatrix} t_1 & t_2 \end{bmatrix}$ indexes the relative timing of the structural change and θ_n is the magnitude of shift. Hansen considers testing $H_0 : \theta_n = 0$ in (8), against a sequence of local alternatives given by

$$\theta_n = \theta + \delta\sigma/\sqrt{n}$$

for fixed δ as $n \to \infty$, and σ^2 is the variance of e_{nt} above. The model can be written as

$$y_t = \begin{cases} x'_{nt}\beta + e_{nt}, & \text{under H}_0 \\ x'_{nt}\beta + x'_{nt}\theta_n I(t \geq T_0) + e_{nt}, & \text{under H}_1 \end{cases} \tag{9}$$

The motivation for formulating the problem in this fashion is to explore the important idea that it is often not known *a priori* whether it is the distribution of β that is broken, or whether the model also contains the feature that the marginal distributions of the independent regressors x_{nt} may also be subject to structural change. In such circumstances, if the null distribution of a test for stability for β (our primary concern) is affected by a structural change in x_{nt} , then a significant test statistic could indicate that there is instability in either the coefficient or in the regressors or a combination of the two, and inference would be distorted.

Under the null hypothesis, the variance estimate is $\widehat{\sigma}^2 = (T-m)^{-1} \sum_{t=1}^{T} \widehat{e_t}^2$ and under the alternative hypothesis, the variance estimate is $\widehat{\sigma_t}^2 = (T - 2m)^{-1} \sum_{t=1}^{T} \widehat{e_{it}}^2$. The break date, $\widehat{t}$, can be estimated as $\widehat{t} = \arg\min \widehat{\sigma_t}^2$.

The standard test for known date of break T_0 of the null against the alternative hypothesis is given by

$$F_t = \frac{(n-m)\widehat{\sigma}^2 - (n-2m)\widehat{\sigma_t}^2}{\widehat{\sigma_t}^2} \tag{10}$$

Since the timing of the break is unknown, Hansen investigates the asymptotic properties (under the assumption of both stationary and non-stationary marginal processes) of the $SupF_T = sup_t F_t$ due to Quandt (1960), and the $ExpF_T = \ln \int exp(F_t/2)dw(t)$and average F test $AveF_T = \int_t F_t dw(t)$, where w is a measure which puts weight $\frac{1}{(t_2-t_1)}$ on each integer t in the interval $[t_1, t_2]$, due to Andrews and Ploberger (1994).

Hansen shows that these tests are not invariant to structural change in the regressors and proposes the use of fixed regressor bootstrap to avoid the size problem of over- or under-rejection of the tests. The bootstrap, which treats the regressors as if they are fixed even when they contain lagged dependent variables, achieve reasonable size properties even in small samples and allows for arbitrary structural change in the regressors, including shifts in mean, polynomial trends, exogenous stochastic trends and the presence of lagged dependent variables and heteroscedastic error processes. This paper thereby contains a major generalisation in terms of considering structural change in systems of equations. The only aspect that is missing from this analysis is the consideration of multiple breaks for structural change which could however be considered within the same framework (and for which we may derive simulation evidence from Banerjee, Lazarova and Urga, 1998, and de Peretti and Urga, 2004).

2.5 Where do the papers in this volume lie

This section of the introduction places the papers in this volume on structural change in perspective within the literature described above. The contributions of these papers lie in their analysis of the precise characterisation of the form of the break and its impact on estimation and inference, in improving generalized method of moments (GMM) estimators for models with structural change and in considering the impact of structural change on forecasting.

2.5.1 Characterisation of the break (Montanes, Olloqui and Calvo; Perron and Zhu; Hillebrand)

In a series of papers, Paul Newbold and his co-authors have argued that the presence of a break in a process which is integrated of order one under the null hypothesis (such as the characterisation presented for the Perron model above) causes routine application of standard Dickey-Fuller tests to lead to some cases of spurious rejections of the unit root hypothesis, if the break occurs early in the series. Furthermore, if the series contain more than one break, only one of which is specifically accounted for in the analysis, a neglected early break can again lead to spurious rejections of the unit root null. This description of the problem appears *inter alia* in Harvey, Leybourne and Newbold (2001), Kim, Leybourne and Newbold (2000), and Leybourne, Mills and Newbold (1998), and thus presents a phenomenon converse to the one typically studied, where if a series is generated by a process that is stationary around a broken trend, conventional Dickey-Fuller tests can have very low power to reject the null of a unit root.

It is an argument that recurs in the work of Lee and Strazicich (2001) who, in common with Harvey *et al.* (2001), deal with the issue of dating the break accurately. Their results show that procedures such as those given by Zivot and Andrews (1992) often identify the break point incorrectly at one period behind the true break point, and this occurs more frequently as the magnitude of the break increases. It is therefore important to take due note, in estimation and testing, of the consequences of including (or not including) a break under the null, and of dating the break accurately. Simple modifications of the test procedure are available to ameliorate both the oversizing and mistiming properties of these tests, but constitute important issues to be borne in mind.

Antonio Montanes, Irene Olloqui and Elena Calvo (in this volume) assess the consequences of mis-specifying the type of break, by analysing the effects of allowing the break to affect the intercept of the trend function, or its slope or both these parameters. The analytical findings contained in their paper show that incorrect specification of the break may even lead the investigator to conclude that the series under investigation possess explosive roots, if the true DGP contains an (unallowed for) change in the slope of the model. *Pierre Perron and Xiaokang Zhu (in this volume)* analyse the consistency, rate of convergence and limiting distributions of parameter estimates in models where the trend function exhibits change at some point in time and looks at the effects of different specifications of the slope shift regressors and the inclusion or exclusion of a level shift regressor.

A related contribution is by *Eric Hillebrand (in this volume)* which looks at the effects of changes not in the parameters of the conditional model but in the volatility process of the errors. By considering a class of generalized autoregressive heteroscedasticity (GARCH) models, for which the occurrence of structural breaks is said to be well known, he shows that unaccounted-for parametric regime changes in GARCH models cause the sum of the autorergressive parameters to converge to one, leading to a finding of spurious persistence and providing a confounding factor in volatility analysis. This result of course finds quite close resonance in the Perron-type results, where the persistence generated by ignoring breaks relates to the estimates of the roots of the time series model itself. Hillebrand notes that a parallel discussion exists in the literature on the connection between structural breaks and the estimation of the parameter of fractional integration, the latter being the topic of the second part of this introduction. Here too, not unexpectedly, neglecting structural breaks also causes and over-estimation of the parameter of fractional integration and leads the reader to believe in the existence of genuine persistence.

2.5.2 GMM estimators (Gagliardini, Trojani and Urga)

Andrews (1993) developed Wald, Lagrange Multiplier and Likelihood ratio tests for structural change based on generalized method of moments (GMM) estimators. *Patrick Gagliardini, Fabio Trojani and Giovanni Urga (in this volume)* propose a new class of robust GMM tests for endogenous structural breaks. The tests are based on supremum, average and exponential functionals derived from robust GMM estimators . A Monte Carlo simulation exercise is conducted to assess the finite sample performance of the new robust tests and to compare their performance with the more classical GMM tests for structural change, including the Andrews (1993) class. It is shown that robust asymptotic tests have higher power and more stable critical values than their more standard equivalents, while in the more realistic case of small sample sizes, bootstrapped robust GMM tests provide higher finite sample efficiency.

2.5.3 Forecasting with structural change (Pesaran and Timmermann)

A vast literature has developed in this topic, including the prominent work of Clements and Hendry, as reported in, for example, Clements and Hendry (1996, 1998) and subsequent papers. More recent work on forecasting with factor models (see for example Stock and Watson, 2002) has taken explicit account of the possibility of structural breaks and has argued that such methods are inherently robustified against the presence of structural change. An important stylised fact is however evident, especially for some of the macroeconomic aggregates commonly forecast (inflation or GDP growth, for example), namely that simple dynamic autoregressive models (suitably robustified if necessary by, for example, intercept corrections, are difficult to beat as forecasting devices. *Hashem Pesaran and Allan Timmermann (in this volume)* consider the small sample properties of forecasts from autoregressive models under structural breaks. The important aims of the paper are to explore the small sample properties of the parameters of AR models in the presence of structural breaks and to ask how much data to use in building a forecasting model - whether, for example, to combine pre-break with post-break data or to use rolling or expanding windows for estimation. The practical recommendations differ, depending on the nature of the data considered. Where the true AR coefficient(s) decline after a break, using pre-break data in the estimation may be efficacious. In general, the authors conclude that there are many scenarios where the inclusion of some pre-break data for purposes of estimation of the parameters of autoregressive models leads to lower biases and lower mean squared forecast errors than if only post-break data is used. Rather surprisingly, this is also true when the post-break window is large, and the post-break data generating process is either highly persistent and/or has a break in the mean or variance. The theoretical arguments are illustrated in an empirical forecasting exercise based on the analysis of forecasts of GDP and industrial production growth, inflation and interest rates for six major economies. This is an area worthy of further investigation, since it is well known that forecasting in the presence of structural breaks is certainly the most realistic but simultaneously a very difficult exercise.

2.6 Some further reflections

The main emphasis in the literature has been on examining, for example, the implications of structural change for tests of the unit root hypothesis, or estimation and inference in processes characterized by multiple breaks where the timing(s) may or may not be assumed to be known *a priori*. A number of additional issues not discussed above are also worthy of mention within the context of this introduction, which while by no means exhausting the vast range of topics explored in the literature are nevertheless important.

A key component of alternative approaches to looking at structural change involves considering a class of models where the change is gradual (or happens continuously). A simple example of such change involves making the distinctions between additive and innovation outlier models (where the former is characterised by discrete change and the latter by continuous), with implications for the treatment of the dynamics and specifications of the models and the tests considered to detect these changes. Perron and Vogelsang (1992) and Vogelsang and Perron (1998) are the most comprehensive treatment of this set of issues. The work of Lin and Terasvirta (1994, 1999) in developing tests for the constancy of parameters, in a scenario where the parameters of the model may change continuously over time, is of importance in this regard, as are the papers by Busetti and Harvey (2001, 2003), who base their tests on a parameterisation of the model where the deterministic components themselves may evolve as random walks under the null hypothesis. Busetti and Harvey (2004) have subsequently developed some of these ideas to look at tests of macroeconomic convergence in panels of data taken, for example, from Maddison (1995, 2001) or the Penn World tables.

Tests for parameter constancy have also been developed for cointegrated vector autoregressive (VAR) models by Hansen and Johansen (1999). Johansen, Mosconi, and Nielsen (2000) extend the VAR cointegration test to the situation with multiple breaks at known times. Graphical procedures based on recursively estimated eigenvalues are used to evaluate the constancy of the long-run relations in the model. They show that both a fluctuation test and a Nyblom (1989) -type LM test for constancy of parameters can be used to test for the constancy of the parameters in the long-run relations. Hansen (2003) generalises the cointegrated VAR model to allow for structural changes, taking the time of the change points and the number of cointegrating or long-run relations as given, and develops likelihood ratio tests for m against $m + k$ structural changes. His methods also allow for the testing of linear restrictions on the short and long-run parameters in the presence of structural changes.

Finally, it is also possible to model the structural change in terms of Markov-switching processes, and to look at the power of tests for unit roots when the data undergo Markov switching. It is not surprising to note that all standard unit root tests, including those robustified to mean and/or trend growth rates have low power against a Markov switching trend.

In summary, the literature on structural change continues to remain an area of much active research. Based on the papers collected in our CD-Rom, a study of the chronological evolution of the papers on structural change shows that approximately 83 papers on structural change were published in the period 1999-2003, compared to 40 during 1994-1998 and 13 in 1989-1994. In our evaluation of the literature above, we have been able to touch upon only some of the important themes, but in the process demonstrated the richness and diversity of the topics studied and the impressive and continuing rate of growth of interest in the field.

3 Long memory processes and fractional cointegration

As discussed earlier, an alternative methodology for explaining persistence in time series has been to look at their short or long memory properties. In applied work, testing whether economic and financial time series exhibit short or long memory properties has become of paramount importance and there is a growing body of theoretical and applied literature dealing with the measurement of dependence structure.

3.1 Long memory - a background

From an empirical perspective, long memory is related to a high degree of persistence of the observed data. This phenomenon was noted for the first time in non-econometric literature. Hurst (1951, 1957), Mandelbrot and Wallis (1968), Mandelbrot (1972) and McLeod and Hipel (1978), among others, provide examples of long memory behaviour in time series of hydrology, geophysics, climatology and other natural sciences[1]. Since then, the phenomenon of slowly

[1] The most famous data set of long memory process is the yearly minimal water levels of the Nile River for the years 1007 to 1206. The persistence of river flows is recognized as

declining autocorrelation has drawn the attention of many researchers from very different areas of research. Beran (1994a) is an excellent introduction to the topic. Baillie (1996), Doukhan, Oppenheim and Taqqu (2003), and Robinson (1994a, 2003) have provide comprehensive surveys on long memory processes and fractional integration while a recent annals Issue of this journal entitled *Long Memory and Nonlinear Time Series,* edited by Davidson and Terasvirta (2002), has dealt with non-stationarity, long memory and non-linearities in time series. The main aim of this section of the introduction is to integrate the literature reviews on long memory mentioned above with the most recent research on the subject, to provide an up-to-date overall picture of the studies on long memory processes and to put in the context the papers in this volume.

The starting point of the literature on fractionally integrated processes has been the fact that many economic and financial time series show evidence of being neither I(0) nor I(1). They show significant autocorrelation up to very long lags, often defined as 'hyperbolic decay'. However, when first differenced, those series appear as being 'overdifferenced'. This feature is typical of long memory processes. The order of integration of such a series is usually denoted 'd' and is a fractional (non integer) number with values less than 0 indicating weak or memory-less processes, the interval $(0, \frac{1}{2})$ for stationary processes wtih long memory, and $(\frac{1}{2}, \infty)$ for long memory non-stationary processes.

In what follows, we provide general definitions of a long memory process (Section 3.2) and (in Section 3.3) related discrete-time theoretical models. In Section 3.4 we present estimation methods distinguishing parameteric from semi-parametric methods and in Section 3.5 we introduce testing procedures for long memory. Section 3.6 begins with a short description of long memory volatility processes. We next discuss the importance of the presence of structural breaks and regime switching as sources of long memory (Section 3.6.1), since this is of considerable interest to the interface between the two sections of our volume, and we consider procedures to test parameter stability in the presence of long memory in time-series regressions (Section 3.6.2).

3.2 Definitions of long memory

Long memory processes are expressed either in the time domain or in the frequency domain. In the time domain, long memory manifests itself as hyperbolically decaying autocorrelation functions. However, since it is the behaviour of autocorrelations at long lags that matters, in small samples it is impossible to report this in a graph. In frequency domain, we get the same information in a form of spectrum which has the advantage of showing all information within the interval $[0, \pi]$. There exist several alternative ways to define long memory and we start by introducing the classical definitions.

In the time domain, a stationary discrete time series process is defined to exhibit long memory or long range dependence if its autocorrelation function ρ_j

Hurst effect, and is commonly found in time series of water level. This phenomenon is not exclusive to the study of hydrology, but many economic or financial time series share the same characteristics.

at lag j satisfies:

$$\lim_{j\to\infty}\frac{\rho_j}{c_\rho j^{-\alpha}}=1 \tag{11}$$

for some constants $0 < c_\rho < \infty$ and $0 < \alpha < 1$. This definition implies that the dependence between successive observations decays slowly as the number of lags tends to infinity. A more general definition has been provided by McLeod and Hipel (1978)

$$\lim_{n\to\infty}\sum_{j=-n}^{n}\left|\,\rho_j\,\right|=\infty \tag{12}$$

with n equal to the number of observations. Both definitions (11) and (12) are consistent with processes that do not have unit roots but whose autocorrelation function does not decay too fast.
Another definition of long memory involves the spectral density. A process with spectral density f is defined to exhibit long memory if

$$\lim_{\lambda\to 0}\frac{f(\lambda)}{c_f\left|\lambda\right|^{-\beta}}=1 \tag{13}$$

for some constants $0 < c_f < \infty$ and $0 < \beta < 1$. This implies that the spectral density is unbounded at low frequencies. Definitions (11) and (13) are not equivalent but connected (Beran, 1994a): if we let $1/2 < H < 1$, then $\alpha = 2 - 2H$ and $\beta = 2H - 1$. H is called the Hurst exponent or Hurst coefficient, as originally defined by Hurst (1951), and it represents the classical parameter characterising long memory. Long memory typically occurs when $1/2 < H < 1$. See also Yong (1974) and Robinson (1995b). Finally, the 'negative memory' or 'antipersistence' case occurs when $-1 < \beta < 0$.

Alternatively, the memory of a process can be expressed in terms of behaviour of its partial sums $S_T = \sum_{t=1}^{T} x_t$. Rosenblatt (1956) defines short range dependence in terms of a process that satisfies strong mixing. Resnick (1987) provides a definition of long memory that includes any process which has an autocovariance function for large k such that $\gamma_k \approx \Xi(k)k^{2H-2}$ where $\Xi(k)$ is any slowly varying function at infinity. Helson and Sarason (1967) show that any process with $H > 0$ and autocovariance function given by the expression above violates the strong mixing condition, and hence is a long memory process. Finally, Diebold and Inoue (2001) note that there is a close connection between the variance-of-partial-sum definition and the spectral and autocorrelation definition of long memory.

Two classes of long memory models have been developed in the literature: continuous-time and discrete-time models. Continuous-time models with long memory have been used to model stochastic volatility (Comte and Renault 1996 and 1998; Comte, Coutin and Renault 2003). Since all the papers in the volume consider only discrete time models, we restrict ourselves to considering only this class in our review.

3.3 Discrete-time long memory models.

A number of models have been proposed to describe the 'long memory' feature of time series, including the 'fractional white noise' model or the concept of 'fractional integration' on ARMA models as below.

Fractional Integrated Processes. The simplest discrete-time long memory model is the fractional white noise (Adenstedt, 1974, Granger, 1980, Granger and Joyeux, 1980, and Hosking, 1981). It is defined as

$$y_t = (1 - L)^{-d} \epsilon_t \tag{14}$$

where L is the lag operator, $E(\epsilon_t) = 0$, $E(\epsilon_t^2) = \sigma^2$, $E(\epsilon_t \epsilon_s) = 0$ for $s \neq t$, and $d = H - 1/2$ is the fractional difference parameter (see Adenstedt, 1974, and Taqqu, 1975). Let $y_t \sim I(d)$. If $d = 0$, $y_t = \epsilon_t$ and the process is serially uncorrelated. If $d > 0$ the process is said to have long memory and is mean square summable. It is also stationary for $d < 1/2$ and invertible for $d > -1/2$.

ARFIMA Processes. A more general class of processes that contains the fractional white noise as a particular case is the Autoregressive Fractionally Integrated Moving Average ($ARFIMA$) model introduced by Granger and Joyeux (1980), Granger (1980, 1981), and Hosking (1981). The $ARFIMA(p, d, q)$ process is defined as:

$$\phi(L)\,(1 - L)^{d}(y_t - \mu) = \theta(L)\,\epsilon_t \tag{15}$$

where $\phi(L)$ and $\theta(L)$ involve autoregressive and moving average coefficients of order p and q respectively and ϵ_t is a white noise process. The roots of $\phi(L)$ and $\theta(L)$ lie outside the unit circle. A fractional white noise process is equivalent to an $ARFIMA(0, d, 0)$ process. $ARFIMA$ processes are covariance stationary for $-1/2 < d < 1/2$, mean reverting for $d < 1$ and weakly correlated for $.d = 0$. For $d > 1/2$ these process have infinite variance. For $d \geq 1/2$ the processes have infinite variance but in the litearture it is more usual to impose initial value conditions (extending that of autoregressive models with unit roots) so that y_t has changing, but finite, variance.

Error Duration Model. As an alternative to the ARFIMA model, Parke (1999) introduces the Error Duration model. The basic mechanism for this model is a sequence of shocks of stochastic magnitude and stochastic duration. The variable observed in a given period is the sum of shocks that survive to that point, and the distribution of the durations of the shocks determines whether or not the process is fractionally integrated. Parke concludes that fractional integration requires that a small percentage of the shocks have long durations.

GARMA Processes. Many economic time series present a persistent periodic behaviour that cannot be captured by classical ARMA or ARFIMA models and several papers have proposed long memory models capable of embodying a possible harmonic component in time series. A generalisation of the ARMA/ARFIMA processes is the so-called Gegenbauer $ARMA$ process ($GARMA$). The $GARMA$ process is defined as:

$$\phi(L)\left(1-2\eta L+L^2\right)^d(y_t-\mu)=\theta(L)\,\epsilon_t \tag{16}$$

where η provides information about the periodic movement in the data and the term $\left(1-2\eta L+L^2\right)^{-d}$ generates Gegenbauer polynomials. For a formal treatment of Gegenbauer polynomials see, for example, Szego (1975). When $\eta=1$, the process reduces to an $ARFIMA(p,2d,q)$, whilst when $\eta=1$ and $d=1/2$ it simplifies to an *ARIMA* process. The process was proposed by Andel (1986) and has been studied, among others, by Gray, Woodward and Zhang (1989, 1994) and Chung (1996a, 1996b). The *GARMA* process is stationary if $d<1/2$ when $|\eta|<1$. When $|\eta|=1$ the process is stationary for $d<1/4$. Furthermore, it is invertible if $d>-1/2$ when $|\eta|<1$or if $d>-1/4$ when $|\eta|=1$.

The paper by *Violetta Dalla and Javier Hidalgo paper (in this volume)* is based on this framework. They propose two different procedures, depending upon whether or not the data exhibits strong cyclical components. One is a Wald type test while the other follows the Lagrange Multiplier principle. Both tests are based on the supremum of a sequence of random variables. In addition, they propose a bootstrap method in the frequency domain as alternative to bootstrap based on time domain.

Fractional cointegration. An important role is played by the concept of fractional cointegration (Granger 1980, 1983). Two time series y_t and x_t, integrated of order d and b respectively, are said to be fractionally cointegrated of order $(d,\, b)$ if the error correction term represented by the linear combination $z_t=y_t-\beta x_t$ is fractionally integrated of order $d-b$, where $0<b\leq d$ and $d>1/2$. It is also possible to have fractional cointegration with $d<1/2$, as in Robinson (1994b). In this case the error correction term is mean reverting and a shock to the system persists for some time but eventually dies out. Long run equilibrium exists amongst the variables even though adjustments to equilibrium may take a long time to realise. Granger (1986) has provided an error correction representation for fractionally cointegrated processes. Specifically, if y_t is an $I(d)$ vector of time series and z_t a set of cointegrating vectors such that the error correction term $z_t=\alpha' y_t$ is $I(d-b)$, the fractionally cointegrated system has an error representation of the form:

$$\Psi(L)(1-L)^d y_t=-\gamma\left(1-(1-L)^b\right)\cdot(1-L)^{d-b}z_t+c(L)\,\epsilon_t \tag{17}$$

where $\Psi(L)$is a polynomial matrix in the lag operator L, $\Psi(0)$ is the identity matrix, $c(L)$ is a finite order polynomial and ϵ_t is a white noise error term. Testing the hypothesis of fractional cointegration requires testing for fractional integration in the error correction term. Dittmann (2004) proposes an alternative error correction model that can be employed to estimate fractionally cointegrated systems in three steps.

There are a number of recent theoretical contributions on fractional cointegration such as Chan and Terrin (1995), Jeganathan (1999), Robinson and Marinucci (2000, 2001), Robinson and Hualde (2003). Most of the existing literature has not allowed for the presence of deterministic trends (Robinson and

Marinucci, 2001) or even if trends are allowed (Robinson and Marinucci, 2000, Chen and Hurvich, 2003), in general the deterministic components "corrupt" the stochastic components. *Peter Robinson and Fabrizio Iacone (in this volume)* evaluate processes that contain additive deterministic trends. They analyse asymptotic properties of three alternative bivariate models, providing also a small Monte Carlo study and an empirical application to testing the PPP hypothesis in three US cities.

In addition to the literature mentioned above, important contributions in the field are the following papers: Dueker and Startz (1998) illustrate a cointegration testing methodology based on joint estimates of the fractional orders of integration of a cointegrating vector and its parent series; Marinucci (2000) uses spectral regression to test for cointegrated time series with long-memory innovations; Davidson (2002) models cointegration in fractionally integrated processes and considers methods for testing the existence of cointegrating relationships using parametric bootstrap; Gil-Alana (2003, 2004) proposes a two-step testing procedure of fractional cointegration in macroeconomic time series, based on the the Robinson (1994c) test; Robinson and Yajima (2002) develop methods of investigating the existence and extent of cointegration in fractionally integrated systems; Velasco (2003) considers consistent estimation of the memory parameters of a nonstationary fractionally cointegrated vector time series; Nielsen (2004) proposes a Lagrange Multiplier test of the null hypothesis of cointegration in fractionally cointegrated models.

Estimation methods and testing procedures to detect long memory are discussed in the following two sections.

3.4 Estimation: parametric and semi-parametric methods

The estimation methods proposed to test for long range dependence can be divided into two classes, 'semi-parametric' estimation and 'parametric' estimation. 'Semi-parametric' methods do not require the modelling of a complete set of the autocovariances, we are only interested in the parameter d. If a complete model is built, such as an $ARFIMA(p, d, q)$, we term the estimation 'parametric'. The main disadvantages of parametric methods are that they are computationally expensive (large number of parameters to estimate) and are subject to mispecification. On the other hand, semi-parametric models consider d as the most important parameter of interest and it is robust to mispecification. However, the semi-parametric estimates are less efficient than well specified parametric counterparts. Robinson (1994a) explains the difference between these two classes.

3.4.1 Parametric Methods

Robinson (1994a) introduces the distinction between an 'efficient Gaussian parametric estimate' and an 'inefficient parametric estimate'. The efficient Gaussian parametric estimate methods are the ones for Gaussian series which have an exact (or well approximated) maximum likelihood. The inefficient parametric estimate may be Gaussian estimates without the Gaussianity assumption, or they may be alternative estimates which are less efficient than the maximum likelihood estimator (MLE) under Gaussianity. Baillie (1996) presents several joint estimation methods of the parameters in the ARFIMA model under the assumption of normality. Sowell (1990, 1992) derives the exact maximum likelihood estimator of the *ARFIMA* process with unconditional normally distributed error term. As with all maximum likelihood estimators, Sowell's estimator can perform poorly if the model is mispecified. Baillie and Chung (1993) develop a conditional sum of squares estimator in the time domain and show that it performs similarly to Sowell's estimator for the $ARFIMA\,(0, d, 0)$ model. Doornick and Ooms (2003, 2004), building on the work by Hosking (1981) and Sowell (1992), show that exact MLE can be efficiently estimated with storage of order n and computation of order n^2, and discuss the necessary efficient computation of the ACF of ARFIMA processes. Their ML approach allows for (break-) regressors in the mean and structural changes in the variances.This is implemented in the Arfima package and therefore also in PcGive (Doornik and Hendry, 2001).

The full MLE is complicated and computationally demanding, but developments in the recent years have facilitated approximations. Procedures for approximation of the log-likelihood have been proposed both in the frequency domain (Whittle 1951; Fox and Taqqu 1986; Beran 1994a; Taqqu, Teverovsky and Willinger 1995) or in the time domain, such as Haslett and Raftery (1989). The Whittle estimator is based on the periodogram. It is defined as the vector η of unknown parameters which minimises the function

$$Q\left(\eta\right) = \int_{-\pi}^{\pi} \frac{I\left(\lambda\right)}{f\left(\lambda;\eta\right)} d\lambda \tag{18}$$

where $I\left(\lambda\right)$ is the periodogram and $f\left(\lambda;\eta\right)$ is the spectral density at frequency λ or its discrete-frequency form

$$Q\left(\eta\right) = \frac{1}{2\pi n} \sum_{j=1}^{n} \frac{I\left(\lambda_j\right)}{f\left(\lambda_j;\eta\right)} \tag{19}$$

where $\lambda_j = \frac{2\pi j}{n}$ are Fourier frequencies. Under an $ARFIMA\,(p, d, q)$ specification, the vector η is an estimate of the autoregressive, moving average and long memory coefficients. Under a fractional white noise specification, η reduces to the parameter d.

To complement the literature discussed so far, it is useful to mention briefly a series of relevant papers on parametric methods: Beran (1994b) proposes a general linear regression and a M-estimators for Gaussian long-memory models, respectively; Beran and Terrin (1996) test for the stability of H; Pai and Ravishanker (1998) use Bayesian analysis to detect changing parameters in ARIMA processes; Bhansali and Kokoszka (2001) present a general overview of estimation of long memory processes; Giraitis *et al.* (2003) introduce a Gaussian estimation of parametric spectral density; Peters and Sibbertsen (2001) extend ML estimators to develop robust tests on fractional cointegration; Andrews and

Lieberman (2002) study the parametric bootstrap for long memory processes; Zaffaroni (2003) presents a Gaussian inference of long-range dependent volatility methods. For more details, interested readers may consult for example Section 1.2 in Robinson (2003).

The above discussion concerns the case $0 < d < \frac{1}{2}$.Velasco and Robinson (2000) extended the Whittle estimation procedure to the case of non stationary ($\frac{1}{2} \leq d < 1$) or antipersistence ($-\frac{1}{2} < d < 0$) processes. For more details, interested readers may consult for example Section 1.5 in Robinson (2003).

3.4.2 Semi-Parametric Methods

Alternative procedures for testing for long memory have been developed. Geweke and Porter-Hudak (1983; GPH henceforth) propose a semiparametric estimator of d in the frequency domain in which the periodogram is first estimated from the series and its logarithm is subsequently regressed on a trigonometric function. However, this test too has several shortcomings. Agiakloglou, Newbold and Wohar (1992), *inter alia*, show that it is biased and inefficient when the error term is $AR(1)$ or $MA(1)$ and in addition does not possess satisfactory asymptotic properties. Robinson (1995a) refines the GPH log-periodogram regression. Using the same notation as GPH, the estimator of the long memory parameter, d , is based on the least-square regression:

$$\log\{I_y(\omega_j)\} = a + b\log(\omega_j) + v_j \tag{20}$$

where $I_y(\omega_j)$ is the periodogram of y_t with spectral ordinates $\omega_1, \omega_2, ..., \omega_m$ and $j = 1, 2, \cdots, m$ and

$$\hat{d} = -\frac{1}{2}\hat{b}$$

The least square estimator $\hat{b}$, which gives $\hat{d}$, is asymptotically normal and the corresponding theoretical standard error is

$$\pi(24m)^{-\frac{1}{2}}$$

An advantage of this version is that it is easier to use for actual computation. The value of the estimator $\hat{d}$ depends on the choice of truncation parameter m. As Diebold and Inoue (2001) point out, the choice of a large value for m, would result in reducing standard error at the expense of inducing bias in the estimator. This is because the relationship that the GPH regression is based on usually holds only for frequencies close to zero. On the other hand, consistency requires that m grows with sample size, but at a slower rate. Diebold and Inoue adapt a 'popular rule of thumb of $m = \sqrt{T}$, where T is the number of observations.

Robinson (1995b) proposes a local Whittle estimator where the objective function is a discrete form of an approximate frequency domain Gaussian likelihood, averaged over a neighbourhood of zero frequency. The neighbourhood grows to infinity but more slowly than the sample size. This estimator is shown to be asymptotically normal and more efficient than previous estimators. Semiparametric estimators also include data differeing and data tapering methods (Velasco 1999; Hurvich and Chen 2000). Phillips and Shimotsu (2004a,b) propose variant of the local Whitle estimation procedure that does not rely on differenicng or tapering and they further extend the range where the estimator of d has standard asymptotics.

Higuchi (1988) introduces an estimator to measure the fractal dimension of d of a non periodic and irregular time series. A test based on Higuchi's estimator can be found in de Peretti (2003). Jensen (1999) proposes an estimator for d that is constructed by means of wavelets. Wavelets are functional transforms that can simultaneously localise a process in time and scale, thus allowing to identify either long run behaviour or short run events more effectively than, for example, Fourier transforms. Jensen's wavelet OLS estimator of d is based on the smooth decay of the autocovariance function of long-memory processes. However, Jensen wavelet decomposition presents noticeable limits as pointed out by Bardet *et al* (2000).

To complement the literature discussed so far, it is important to mention Robinson (1994b) and Lobato and Robinson (1996), who propose an average periodogram estimation for long memory; Lobato (1999) analyses a two-step estimator of a long memory vector process; Bollerslev and Wright (2000) perform a semiparametric estimation of long memory volatility dependencies; Wright (2000) develops log-periodogram estimators with conditional heavy tails; Henry (2001) introduces a periodogram spectral estimation for the case of long memory conditional heteroscedasticity; Arteche (2002) deals with the problem of seasonal and cyclical long memory developing semiparametric robust tests; Christensen and Nielsen (2002) perform a semiparametric analysis of fractional cointegration. Whitcher (2004) provides an up-to-date account of estimating multifactor GARMA models using (semi-) parametric wavelet methods for long memory processes. Finally, Robinson (2003, Section 1.2) offers a detailed discussion on the topic.

3.5 Testing for Long Memory

Many tests developed for long memory are based on the estimation of the Hurst exponent. The statistic very much used over the years is the rescaled range (R/S henceforth) statistic introduced by Hurst (1951) and used, for example, by Mandelbrot (1972, 1975). The R/S statistic is the range of partial sums of deviations from the mean, rescaled by its standard deviation. Formally:

$$\bar{Q}_n = \frac{R_n}{s_n} = \frac{\max\limits_{1\le k\le n} \sum\limits_{j=1}^{k} \left(X_j - \bar{X}_n\right) - \min\limits_{1\le k\le n} \sum\limits_{j=1}^{k} \left(X_j - \bar{X}_n\right)}{\sqrt{\frac{1}{n} \sum\limits_{j=1}^{n} \left(X_j - \bar{X}_n\right)^2}} \tag{21}$$

where $\bar{X}_n = \frac{1}{n}\sum_{j=1}^{n} X_j$ is the sample mean and s_n is the sample standard deviation. However, the R/S statistic lacks robustness in the presence of short memory and heteroskedasticity, as well documented in Mandelbrot, 1972, 1975; Mandelbrot and Wallis, 1968; Davies and Harte, 1987; Aydogan and Booth, 1988; Lo, 1991; among others. Lo (1991) proposes a modified rescaled range (MR/S henceforth) test, defined as:

$$Q_n = \frac{R_n}{\hat{\sigma}_n(q)} \tag{22}$$

where

$$\hat{\sigma}_n^2(q) = \hat{c}_0 + 2\sum_{j=1}^{q} w_j(q)\,\hat{c}_j \tag{23}$$

$\hat{c}_j$is the jth order sample autocovariance of X_j and $w_j(q) = 1 - \frac{j}{q+1}$, $q < n$, are the Bartlett window weights. Lo shows that this test is consistent against a general class of long range dependent stationary Gaussian alternatives. However, a practical difficulty arises in the choice of q and in distinguishing between short range and long range dependencies. Hauser and Reschenhofer (1995) apply the R/S test statistic of Davies and Harte (1987) to estimate d.

Though tests based on R/S statistics are easy to use, they are very much *ad hoc* and involve non standard asymptotics. Alternatively, tests can be constructed on the basis of parametric or semi-parametric MLE of d (see for instance Lobato and Robinson 1998). These tests produce standard asymptotics so that classical rules for inference are available. In case of correct specification of parametric models these tests display optimality properties.

We close this section by highlighting the rescaled variance statistics (Giraitis et al, 2003) and the t_n automatic bandwidth selection statistic proposed by Lobato and Robinson (1998). A note of interest is that unit root tests can be used to test for long memory. Lee and Schmidt (1996) show that the Kwiatkowsky, Phillips, Schmidt and Shin (1992) test, originally developed to test for an $I(0)$ null hypothesis against an $I(1)$ alternative is consistent against an $I(d)$ alternative and it can therefore be used to distinguish short memory from long memory stationary processes. A number of studies (Diebold and Rudebusch, 1991 and Hassler and Wolters, 1994, amongst others) claim that the Dickey-Fuller, Augmented Dickey-Fuller and Phillips-Perron unit root tests are consistent against $I(d)$ alternatives, although performing relatively poorly in distinguishing between the $I(1)$ null hypothesis and the $I(d)$ alternative. Finally, Dittmann (2000) reports an extensive Monte Carlo study comparing size and power of six residual-based tests (Lobato-Robinson LM test, the Geweke-Porter-Hudak test, two Phillips-Perron-type tests, the Dickey-Fuller test and the modified rescaled range test) of the hypothesis of no cointegration against the alternative of fractional cointegration. Finally, a new test for fractionally integrated processes is proposed by Dolado, Gonzalo and Mayoral (2002), where the testing procedure is in the time domain and extends the well-known Dickey-Fuller approach.

3.6 Long memory processes, breaks and regime switching

Long memory may be a genuine feature of time series. Moreover, long memory may also occur when several processes are aggregated (Robinson 1978; Granger 1980) or because of the presence of 'structural breaks' and 'regime switching' in the series. In this section of our review, we will discuss how recent literature has dealt with the presence of 'structural breaks' and 'regime switching' in studying long memory processes in the data. Granger and Terasvirta (1999) show how the estimate of long memory parameter very much depends on the number of regime switches and where they occur in the sample. Gourieroux and Jasiak (2001) study how processes with infrequent regime switching may generate a long memory effect in the autocorrelation function. In this case fractional I(d) model for economic or financial analysis may lead to spurious results. Granger and Hyung (2004) notice that a linear process with breaks can mimic the properties of long memory.

Diebold and Inoue (2001) show that stochastic regime switching is easily confused with long memory. Using a definition of long memory that involves the rate of growth of variances of partial sums $S_T = \sum_{t=1}^{T} x_t$,

$$\mathrm{var}(S_T) = O(T^{2d+1}) \tag{24}$$

the authors explain that there is a close connection between variance-of-partial-sum definition and the spectral and autocorrelation definitions of long memory. The spectral density at frequency zero is the limit of $(\frac{1}{T})S_T$, a covariance stationary process has long memory in the sense of spectrum if and only if it has long memory for some $d > 0$ in the variance-of-partial-sum sense. They employ the variance-of-partial-sum definition of long memory (24). They find evidence of long memory using 'a mixture model with constant break size, and break probability dropping with T', 'the stochastic permanent break model' (Engle and Smith, 1999) and 'the Markov-switching model'[2].

3.6.1 Long Memory and Structural Change in Variance

We may also have long memory volatility process, i.e. long memory in variance. Robinson (1991), Breidt, Crato and de Lima (1998), Giraitis and Robinson (2001), Robinson and Zaffaroni (1997, 1998), just to cite a few papers, have generated a substantial body of literature. Andreou and Ghysels (2002) and Franses, van der Leij and Paap (2002) evaluate how measures of volatility display structural breaks and/or long memory. Mikosch and Starica (2004) evaluate the possible long-range dependence in volatility in presence of integrated GARCH and structural breaks (see also the paper by *Eric Hillebrand (in this volume)*).

From a different viewpoint, Liu (2000) studies the case of high persistence/long memory in volatility caused by regime switching. The main feature of his model is that long memory is generated as a heavy-tailed duration distribution. He

[2] Leipus and Viano (2003) showed that the behaviour of partial sums of the mixture model of Diebold and Inoue differs from the typical convergence to fractional Brownian motion.

proposes a regime switching stochastic volatility (RSSV) model consisting of mean equation

$$y_t = \mu_0 + \mu_1 y_{t-1} + \mu_2 y_{t-2} + u_t \tag{25}$$

and variance equation

$$u_t = e^{w_t} \epsilon_t \tag{26}$$

where w_t is the regime switching variable, with the probability law of duration of regime governed by some heavy-tail distribution L_T. The regime switching variable taking values according to some distribution L_w, ϵ_t is distributed according to law L_ϵ. u_t, e^{w_t} and ϵ_t, are independent. The variance equation is a special feature of RSSV model, and it is also assumed that,

$$w \sim N(\mu_w, \sigma_w^2)$$
$$\epsilon \sim N(0,1)$$
$$\text{Prob}(T_k \geq t) = (1 + ct)^{-\alpha}$$

with $0 < \alpha < \infty$, where t is an integer greater than 0. T_k denotes the duration of the regime and all the quantities above are independent. In duration distribution, α frames the tail behaviour of the distribution, and c, a positive constant, controls the scale. When $1 < \alpha < 2$, the resulting time series exhibits the long memory property.

The results show that the model fits the dynamics of the S&P composite return series extremely well and the estimated tail index is highly significant.The heavy-tailed regime switching in the volatility of the RSSV model also provides enough structure to yield a useful volatility prediction. Finally, Liu interprets the observed persistence as arising from volatility regime switching, showing that when the duration of the regime has a heavy-tail distribution, long memory process does not exist.

In the same line of research, *Remigijus Leipus, Vygantas Paulauskas and Donatas Surgailis (in this volume)* evaluate the long memory properties of an AR(1) with drift ($X_t = \mu_t + a_t X_{t-1} + \sigma_t \varepsilon_t$) with renewal regime affecting the shift/levels (μ_t), the slope coefficient (a_t) and/or the volatility (σ_t) of the process. They expand a companion work (Leipus and Surgailis, 2003) and compare their regime switching volatility model with that of Liu's (2000).

A partial sample of recent papers reporting empirical applications include Barkoulas, Baum and Caglayan (1999) who test for fractional integration and mean shifts in real exchange rates in 12 countries; Bos, Franses and Ooms (1999) who evaluate the implication for long memory neglecting mean shifts in G7 inflation rates; and Hyung and Franses (2001, 2002) who deal with long memory and breaks in US inflation rates.

3.6.2 Long memory and breaks in time series

Finally, we consider the case where the breaks (structural change) and/or long memory are genuine features of time series. A series of contributions have proposed testing procedures for evaluating changing parameters in presence of long memory. Kuan and Hsu (1998) analyze the least-squares estimator of the change point for fractionally integrated processes with fractional differencing parameter $-0.5 < d < 0.5$. When there is a one-time change, the authors show that the least-squares estimator is consistent and that the rate of convergence depends on d. When there is no change, they find that the least-squares estimator converges in probability to the set $\{0, 1\}$ for$-0.5 < d < 0$ but is likely to suggest a spurious change for $0 < d < 0.5$. Simulations are also used to illustrate the asymptotic analysis. Wright (1998) derives the null limiting distribution of the sup-Wald and CUSUM test for structural stability with an unknown potential break date in a polynomial regression model where the errors are $I(d)$, $-0.5 < d < 0.5$. For $d > 0$, both tests diverge under the null hypothesis so that the asymptotic size of either test is unity. For $d < 0$, both tests converge to zero under the null so that the asymptotic size of either test is zero. Lavielle and Moulines (2000) use penalized least-squares to estimate the number of known change-points in presence of long memory in the error process. Kramer and Sibbertsen (2002) derive the limiting null distributions of the standard and OLS-based CUSUM- tests for a structural change of the coefficients of a linear regression model in the context of long-memory disturbances. They show that both tests have different performance in presence of long-memory as compared to short memory, and that long memory is easily mistaken for structural change when standard critical values are employed. Ohanissian, Russell, and Tsay (2003) develop a simple test to distinguish between spurious long memory due to breaks and true long memory. Hidalgo and Robinson (1996) propose tests for a change in parameter values at a known time point in linear regression models with long-memory errors. The tests are derived in case of certain nonstochastic and stochastic regressors, and are given large-sample justification. A small Monte Carlo study of finite-sample behaviour is included. *Stepana Lazarova (in this volume)* introduces a new test for structural change in time series regression where both regressors and residuals exhibit long range dependence.The author proposes a bootstrap procedure to approximate the distribution of the test and reports a small Monte Carlo exercise to validate the robustness of the proposed method.

3.7 Where do the papers in this volume lie

The literature review on long memory presented in this section has put in perspective recent research on the subject and has put in the context the papers in this volume. The contributions of these papers lie in their analysis of bootstrap methods to test for the presence of strong cycles, of fractional contegration in presence of deterministic trends, in evaluating long memory in renewal regime switching schemes and in testing for structural change in presence of long range dependence.

3.7.1 Cyclical data and dependence (Dalla and Hidalgo)

This paper proposes a simple test for the hypothesis of strong cycles and, as a by-product, a test for weak dependence for linear processes. The authors show that the limit distribution of the test coincides with the distribution of the maximum of a (semi-)Gaussian process $G(\tau)$, $\tau \in [0,1]$. Because the covariance structure of $G(\tau)$ is a complicated function of τ and model dependent, to obtain the critical values (if possible) of $max_{\tau\in[0,1]}G(\tau)$ may be difficult. For this reason the authors propose a bootstrap scheme in the frequency domain to circumvent the problem of obtaining (asymptotically) valid critical values. The proposed bootstrap can be regarded as an alternative procedure to existing bootstrap methods in the time domain such as the residual-based bootstrap. Finally, they illustrate the performance of the bootstrap test by a small Monte Carlo experiment and an empirical example.

3.7.2 Fractional cointegration and deterministic trends (Robinson and Iacone)

The authors consider a cointegrated system generated by processes that may be fractionally integrated, and by additive polynomial or more general deterministic trends. In view of the consequent competition between stochastic and deterministic trends, they consider statistical inference on the cointegrating vector and develop relevant asymptotic theory, including the situation where fractional orders of integration are unknown

3.7.3 Long memory and regime switching (Leipus, Paulauskas and Surgailis)

This contribution discusses long memory properties and large sample behavior of partial sums in a general renewal regime switching scheme. The linear model $X_t = \mu_t + a_t X_{t-1} + \sigma_t \varepsilon_t$ with renewal switching in levels, slope or volatility and generally (possible heavy-tailed) i.i.d. innovations is discussed in detail. Conditions on the tail behavior of interrenewal distribution and the tail index $\alpha \in (0,2]$ of ε_t are obtained, in order that the partial sums of X_t are asymptotically $\lambda - stable$ with index $\lambda < \alpha$.

3.7.4 Long memory and breaks in time series (Lazarova)

The paper considers tests for structural change in time series regression models where both regressors and residuals may exhibit long range dependence. The limiting distribution of the test statistic depends on unknown parameters. While the unknown parameters can be consistently estimated and asymptotic critical values obtained by simulation, the paper proposes an alternative approach of approximating the distribution of the test statistic by a bootstrap procedure. The asymptotic validity of bootstrap is shown and the performance of the testing procedure is examined in a simple Monte Carlo experiment

3.8 Final remarks

It has been a pleasure to collaborate with the contributors to this volume and with our referees in preparing this special annals issue. No doubt,the topic covered by the papers are in areas of much active research and the topic on which a great deal of important methodological advancements have been made for econometric analysis over the last two decades. The bibliographic material collected in our CD-Rom embodies approximately 140 papers on structural breaks, defined fairly narrowly, and 180 papers on testing and estimating the dependence structure of economic and financial data, combining both theoretical and applied work. In addition to studies using using macroeconomic data, recent empirical work uses foreign exchange and stock market data, interest rates and bond data. Further challenges lie ahead and we foresee important issues to explore mainly in the intersection of the two important characteristics of the data, namely the presence at the same time of structural changes and long range dependence.

Acknowledgements

We are indebted to Juan Cajigas for collecting the preparatory bibliographic material for this editorial. We also wish to thank J. Cajigas, L. della Ratta, C. de Peretti, J. Doornik, C. Granger, S. Lazarova, R. Leipus, M. Ooms, P. Perron, P. Robinson and P. Zaffaroni for lengthy discussions and very helpful comments on various drafts of the paper. Responsibility for any remaining errors rests with us. This work was initiated under ESRC project N. R000238145 and has also benefited from funding support from the Research Council of the European University Institute, Florence for the first named author and from the Cass Business School for the second named author. Their support is gratefully acknowledged.

We also wish to thank the Editor, Cheng Hsiao, and an Associate Editor for their support and encouragement throughout the project.

14 papers were submitted for this special volume and 9 were accepted for publication. Each paper was refereed anonymously by at least three referees, and an Associate Editor was one the referees for the contribution from the co-editor. The full list of the referees involved in this volume is given below. To them we extend our special thanks for their cooperation, valuable comments and helpful suggestions. Without them we would not have been able to bring this volume to press.

J. ARTECHE (Departamento de Economía Aplicada III, Universidad del Pais Vasco, Bilbao, Spain)

J. BAI (Department of Economics, New York University, USA)

J.BERAN (Fachbereich für Mathematik und Statistik. Universität Konstanz, Germany)

J. BREITUNG (Bonn Graduate School of Economics, University of Bonn, Germany)

J. BREUER (Department of Economics, Moore School of Business, University of South Carolina, USA)

F. BUSETTI (Research Department, Bank of Italy)

M. CARRASCO (Economics Department, University of Rochester, USA)

M. CLEMENTS (Department of Economics, University of Warwick, U.K.)

M. DACOROGNA (Converium Limited, Zurich, Switzerland)

J. E.H. DAVIDSON Department of Economics, University of Exeter, U.K.)

F.H. FRANSES (Econometric Institute, Erasmus University, Rotterdam, The Netherlands)

A. FUERTES (Faculty of Finance, Cass Business School, U.K.)

J. GEWEKE (Department of Economics and Statistics, University of Iowa, USA)

R. GENCAY (Department of Economics, University of Windsor, Ontario, Canada)

E. GHYSELS (Department of Economics, University of North Carolina, USA)

L. GIRAITIS (Department of Mathematics and Economics, University of York, U.K.)

References

Adelman, I., 1965. Long cycles - fact or artifact? American Economic Review 55, 444-463.

Andel, J., 1986. Long memory time series models. Kybernetika 22, 105-123.

Andenstedt, R.K., 1974. On large sample estimation for the mean of a stationary random sequence. Annals of Mathematical Statistics 2, 1095-1107.

Agiakloglou, C., Newbold, P., Wohar, M., 1992. Bias in an estimator of the fractional difference parameter. Journal of Time Series Analysis 14, 235-246.

Andreou, E., Ghysels, E., 2002. Detecting multiple breaks in financial market volatility dynamics. Journal of Applied Econometrics 17, 579-600.

Andrews, D. W. K., 1993. Tests for parameter instability and structural change with unknown change point. Econometrica 61, 821-856.

Andrews, D. W. K., Lieberman, O. 2002. Higher-order improvements of the parametric bootstrap for long-memory Gaussian processes. Cowles Foundation, Discussion Paper No. 1378, Yale University.

Andrews, D. W. K., Ploberger, W., 1994. Optimal tests when a nuisance parameter is present only under the alternative. Econometrica 62, 1383-1414.

Arteche, J., 2002. Semiparametric robust tests on seasonal or cyclical long memory time series. Journal of Time Series Analysis 23, 251-285.

Aydogan, K., Booth, G. G., 1988. Are there long cycles in common stock returns? Southern Economic Journal 55, 141-149.

Bai, J., 1999. Likelihood ratio tests for multiple structural change. Journal of Econometrics 91, 299-323.

Bai, J, Perron, P., 1998. Estimating and testing linear models with multipple structural changes. Econometrica 66, 47-78.

Bai, J., Perron, P., 2003a. Computation and analysis of multiple structural change models. Journal of Applied Econometrics 18, 1-22.

Bai, J., Perron, P., 2003b. Critical values for multiple structural change tests. The Econometrics Journal 6, 72-78.

Baillie, R. T., 1996. Long memory processes and fractional integration in econometrics. Journal of Econometrics 73, 5-59

Baillie, R.T., Chung, C.F., 1993. Small sample bias in conditional sum of squares estimators of fractionally integrated ARMA models. Empirical Economics 18, 791-806.

Banerjee, A., Lazarova, S., Urga, G., 1998. Bootstrapping sequential tests for multiple structural breaks, Discussion Paper No.17-98, London Business School.

Banerjee, A, Lumsdaine, R.L., Stock, J.H., 1992. Recursive and sequential tests of the unit root and trend break hypotheses: theory and international evidence. Journal of Business and Economic Statistics 10, 271-287.

Bardet, J. M., Lang, G., Moulines, E., Soulier, P., 2000. Wavelet estimator of long-range dependence processes. Statistical Inference for Stochastic Processes 3, 85-99.

Barkoulas, J.T., Baum, C.B., Caglayan, M.., 1999. Long memory or structural breaks: Can either explain nonstationary real exchange rates under the current float? Journal of International Financial Markets, Institutions and Money 9, 359-376.

Beran, J., 1994a. Statistics for long memory processes. Chapman and Hall, New York.

Beran, J., 1994b. On a class of M-estimators for Gaussian long-memory models. Biometrika 81, 755-766.

Beran, J., Terrin, N., 1996. Testing for a change of the long-memory parameter. Biometrika 83, 627-638.

Bhansali, R. J., Kokoszka, P.S., 2001. Estimation of the long-memory parameters: a review of recent developments and an extension. In I.V. Basawa, Heyde, C. C., and Taylor, R.L. (eds), Proceedings of the Symposium on Inference for Stochastic Processes, 125-150

Bollerslev, T., Wright, J.W., 2000. Semiparametric estimation of long-memory volatility dependencies: the role of high-frequency data. Journal of Econometrics 98, 81-106.

Bos, C., Franses, S. P. H., Ooms, M., 1999. Long-memory and level shifts: re-analyzing inflation rates. Empirical Economics 24, 427-449.

Breidt, F. J., Crato, N., de Lima, P., 1998. The detection and estimation of long memory in stochastic volatility. The Journal of Econometrics 83, 325-248.

Busetti, F. and Harvey, A., 2001. Testing for the presence of a random walk in series with structural breaks. Journal of Time Series Analysis 22, 127-150.

Busetti, F. and Harvey, A., 2003. Further comments on stationarity tests in series with structural breaks, Journal of Time Series Analysis 24, 137-140.

Chan, N. H., Terrin, N., 1995. Inference for unstable long-memory processes with applications to fractional unit root autoregression. Annals of Statistics 23, 1662-1683.

Christensen, B. J., Nielsen, M. O., 2002. Semiparametric analysis of stationary fractional cointegration and the implied-realized volatility relation. University of Aarhus Working Paper No. 2001-4.

Chen, W. W., Hurvich, C. M., 2003. Estimating fractional cointegration in the presence of polynomial trends. Journal of Econometrics 117, 95-121.

Chu, C. S., Stinchcombe, M., White, H., 1996. Monitoring structural change. Econometrica 64, 1945-1065.

Chung, C., 1996a. Estimating a generalized long memory process. Journal of Econometrics 73, 237-259.

Chung, C., 1996b. A generalized fractionally integrated autoregressive moving-average process. Journal of Time Series Analysis 17, 111-140.

Clements, M. P., Hendry, D. F., 1996. Intercept corrections and structural change. Journal of Applied Econometrics11, 475-494.

Clements, M.P. and Hendry, D.F., 1998. Forecasting Economic Time Series. Cambridge University Press, Cambridge.

Comte, F. and E. Renault, G.,1996. Long memory continuous time models. Journal of Econometrics 73, 101-149.

Comte, F. and E. Renault, G.,1998. Long memory in continuous-time stochastic volatility models. Mathematical Finance 8, 291-323.

Comte, F., Coutin, L. and E. Renault, E., 2003. Affine Fractional Stochastic Volatility Models with Application to Option Pricing', Discussion paper n. 2003-10, Université de Montreal..

Dalla, V. , Hidalgo, J., 2004. A parametric bootstrap test for cycles, this issue of the Journal.

Davidson, J.,2002. A model of fractional cointegration, and tests for cointegration using the bootstrap. Journal of Econometrics 110,187-212.

Davidson, J., Terasvirta, T. T., 2002. (eds) Long memory and nonlinear time series. Journal of Econometrics 110, Issue 2, 105-437.

Davies, R. B., Harte, D. S., 1987. Tests for the Hurst effect. Biometrika 74, 95-102.

dePeretti, C., 2003. Bilateral bootstrap tests for long memory: an application to the silver market. Computational Economics 22, 187-212.

de Peretti, C., Urga, G., 2004. Stopping tests in the sequential estimation for multiple structural breaks. Working Paper CEA-06-2004, Centre for Econometric Analysis, Cass Business School, London.

Diebold, F. X., Inoue, A., 2001. Long memory and regime switching'. Journal of Econometrics 105, 131-159.

Diebold, F. X., Rudebusch, G. D., 1991. On the power of Dickey-Fuller tests against fractional alternatives. Economics Letters 35, 155-160.

Dittmann, I., 2000. Residual-based tests for fractional cointegration: a Monte Carlo study. Journal of Time Series Analysis 21, 614-647.

Dittmann, I., 2004. Error correction models for fractionally cointegrated time series. Journal of Time Series Analysis 25, 27-32.

Dolado, J. J., Gonzalo, J. , Mayoral, L., 2002. A fractional Dickey-Fuller tests for unit toots. Econometrica 70, 1963-2006.

Doornik, J. A., Hendry, D. F., 2001. Econometric Modelling Using PcGive. (Volume III.) Timberlake Consultants Press, London.

Doornik, J.A., Ooms, M., 2003. Computational aspects of maximum likelihood estimation of autoregressive fractionally integrated moving average models. Computational Statistics and Data Analysis 41, 333-348.

Doornik, J. A., Ooms, M. , 2004. Inference and forecasting for ARFIMA models with an application to US and UK inflation. Studies in Nonlinear Dynamics and Econometrics 8, No. 2, Article 14. http://www.bepress.com/snde/vol8/iss2/art14.

Doukhan, P., Oppenheim, G. and Taqqu, M.S. 2003. (eds) Theory and applications of long-range dependence. Birkhauser, Berlin.

Dueker, M., Startz, R., 1998. Maximum-likelihood estimation of fractional cointegration with an application to U.S.and Canadian bond rates. Review of Economics and Statistics 83, 420-426.

Dufour, J-M., Ghysels, E., 1996. (eds) Recent developments in the econometrics of structural change. Journal of Econometrics 70, Issue 1, 1-316.

Engle, R.F., Granger, C.W.J., 1987. Co-integration and error correction: representation, estimation and testing. Econometrica 55, 251-276.

Engle, R. F., Smith, A. D., 1999. Stochastic permanent breaks. Review of Economics and Statistics 81, 553-574.

Fox, R., Taqqu, M. S., 1986. Large sample properties of parameter estimates for strongly dependent stationary gaussian time series. Annals of Statistics 14, 517-532.

Franses, P. H., van der Leij, M., Paap, R., 2002. Modelling and forecasting level shifts in absolute returns. Journal of Applied Econometrics 17, 601-616.

Gagliardini, P., Trojani, F., Urga, G., 2004, Robust GMM tests for structural breaks, this issue of Journal.

Geweke, J., Porter-Hudak, S., 1983. The estimation and application of long memory time series models. Journal of Time Series Analysis 4, 221-238.

Gil-Alana, L. A., 2003. Testing of fractional cointegration in macroeconomic time series. Oxford Bulletin of Economics and Statistics 65, 517-529.

Gil-Alanal L. A., 2004. A joint test of fractional integration and structural breaks at a known period of time. Journal of Times Series Analysis 25, 681-700.

Giraitis, L., Robinson, P. M., 2001. Whittle estimation of ARCH models. Econometric Theory 17, 608-631.

Giraitis, L., Kokoszka, P., Leipus, R., Teyssiere, G., 2003. Rescaled variance and related tests for long memory in volatility and levels. Journal of Econometrics 112, 265-294.

Gourieroux, C., Jasiak, J., 2001. Memory and infrequent breaks. Economics Letters 70, 29-41.

Granger, C. W. J., 1980. Long memory relationships and the aggregation of dynamic models. Journal of Econometrics 14, 227-238.

Granger, C. W. J., 1981. Some properties of time series data and their use in econometric specification. Journal of Econometrics 16, 121-130.

Granger, C. W. J., 1983. Cointegrated variables and error correction models. Unpublished Manuscript, University of California, San Diego.

Granger, C. W. J., 1986. Developments in the study of cointegrated economic variables. Oxford Bulletin of Economics and Statistics 48, 213-228.

Granger, C. W. J. , 2004. The past and future of empirical finance. Some personal comments (Foreword), this issue of the Journal.

Granger, C. W. J., Hyung, N., 2004. Occasional structural breaks and long memory with an application to the S&P 500 absolute stock returns. Journal of Empirical Finance 11, 399-421.

Granger, C. W. J., Joyeux, R. , 1980. An introduction too long-memory time series models and fractional differencing. Journal of Time Series Analysis 1, 15-29.

Granger, C. W. J., Terasvirta, T., 1999. A simple nonlinear time series model with misleading linear properties. Economics Letters 62, 161-165.

Gray, H. L., Woodward, W., Zhang, N., 1989. On generalized fractional processes. Journal of Time Series Analysis 10, 233-257.

Gray, H. L., Woodward, W., Zhang, N., 1994. On generalized fractional processes: a correction. Journal of Time Series Analysis 15, 561-562.

Hall, R. E., 1979. Stochastic implication of the life cycle-permanent income hypothesis: theory and evidence. Journal of Political Economy 86, 971-987.

Hannan, E.J., Poskitt, D.S., 1988. Unit canonical correlations between future and past. The Annals of Statistics 16, 784-790.

Hansen, B.E., 1992. Tests for parameter instability in regressions with I(1) processes. Journal of Business and Economic Statistics 10, 321-335.

Hansen, B.E., 2000. Testing for structural change in conditional models. Journal of Econometrics 97, 93-115.

Hansen, P.R., 2003. Structural breaks in the cointegrated vector autoregressive model. Journal of Econometrics 114, 261-295.

Hansen, H., Johansen, S., 1999. Some tests for parameter constancy in cointegrated VAR-models. Econometrics Journal 2, 306-333.

Harvey, D.I., Leybourne S.J., Newbold, P. 2001. Innovational outlier unit root tests with an endogenously determined break in level. Oxford Bulletin of Economics and Statistics 63, 559-575.

Haslett, J., Raftery, A. E., 1989. Space-time modelling with long-memory dependence: assessing Ireland's wind power resource. Journal of the Royal Statistical Society (Series C) 38, 1-50.

Hassler, U., Wolters, J., 1994. On the power of unit roots against fractional alternatives. Economics Letters 45, 1-5.

Hauser, M. A., Reschenhofer, E., 1995. Estimation of the fractionally differencing parameter with the R/S method. Computational Statistics and Data Analysis 20, 569-579.

Helson, J., Sarason, Y., 1967. Past and future. Mathematica Scandanavia 21, 5-16.

Henry, M., 2001. Averaged periodogram spectral estimation with long memory conditional heteroscedasticity. Journal of Time Series Analysis, 22, 431-459.

Hidalgo, J., Robinson, P.M., 1996. Testing for structural change in a long-memory environment. Journal of Econometrics 70, 159-174.

Higuchi, T., 1988. Approach to an irregular time series on the basis of the fractal theory. Physica D 31, 277-283.

Hillebrand, E., 2004. Neglecting parameter changes in GARCH models, this issue of the Journal.

Hosking, J. R. M., 1981. Fractional differencing. Biometrika 68, pp. 165-176.

Hurst, H. E., 1951. Long-term storage capacity of reservoirs. Transactions of the American Society of Civil Engineers 116, 770-799.

Hurst, H. E., 1957. A suggested statistical model of some time series that occur in nature. Nature 180, 494.

Hurvich, C., M., Chen, W. W., 2000. An efficient taper for potentially overdifferenced long-memory time series. Journal of Time Series Analysis 21, 155-180.

Hyung, N., Franses, P. H., 2001. Structural breaks and long memory in US inflation rates: do they matters for forecasting? Erasmus University Rotterdam, Econometric Institute Research Report EI2001-13.

Hyung, N., Franses, P. H., 2002. Inflation rates long-memory, level, shifts, or both? Econometric Institute Report 2002-08, Erasmus University.

Jeganathan, P. , 1999. On asymptotic inference in cointegration time series with fractionally integrated errors. Econometric Theory 15, 583-621.

Jensen, M. J., 1999. Using wavelets to obtain a consistent ordinary least squares estimator of the long-memory parameter. Journal of Forecasting 18, 17-32.

Johansen, S., Mosconi, R., Nielsen, B., 2000. Cointegration analysis in the presence of structural breaks in the deterministic trend. The Econometrics Journal 3, 216-249.

Kim, T-H., Leybourne S.J., Newbold,P., 2000. Spurious rejections by Perron tests in the presence of a break. Oxford Bulletin of Economics and Statistics 62, 433-444.

Kramer, W., Sibbertsen, P. 2002. Testing for structural change in the presence of long memory. International Journal of Business and Economics 1, 235-242.

Kuan, C.M., Hsu, C.C., 1998. Change point estimation of fractionally Integrated Processes. Journal of Time Series Analysis 19, 693-708.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P. and Shin, Y., 1992. Testing the null hypothesis of stationarity against the alternative of a unit root: how sure are we that economic time series have a unit root?. Journal of Econometrics 54, 159-178

Lavielle, M., Moulines, E., 2000. Least-squares estimation of an unkown number of shifts in a time series. Journal of Time Series Analysis 21, 33-59.

Lazarova,S., 2004. Testing for structural change in regression with long memory processes, this issue of the Journal.

Lee, D., Schmidt, P. 1996. On the power of the KPSS Test of stationarity against fractionally integrated alternatives. Journal of Econometrics 73, 285-302.

Lee, J., Strazicich, M. 2001. Break point estimation and spurious rejections with endogenous unit root tests. Oxford Bulletin of Economics and Statistics 62, 535-558.

Leipus, R., Paulauskas, V., Surgailis, D. , 2004. Renewal regime switching and stable limit laws, this issue of the Journal.

Leipus, R., Surgailis, D. , 2003. Random coefficient autoregression, regime switching and long memory. Advances in Applied Probability 35, 737-754.

Leipus, R., Viano, M. C., 2003. Long memory and stochastic trend. Statistics and Probability Letters 61, 177-190.

Leybourne, S.J., Mills, T.C., Newbold, P., 1998. Spurious rejections by Dickey-Fuller tests in the presence of a break under the null. Journal of Econometrics 87, 191-203.

Lin, C.-F., Terasvirta, T. 1994. Testing the constancy of regression parameters against continuous structural change. Journal of Econometrics 62, 211-228.

Lin, C.-F., Terasvirta, T. 1999. Testing parameter constancy in linear models against stochastic stationary parameters. Journal of Econometrics 90, 193-213.

Liu, M., 2000. Modelling Long Memory in Stock Market Volatility. Journal of Econometrics 99, 139-171.

Lo, A. W., 1991. Long term memory in stock market prices. Econometrica 59, 1279-1313.

Lobato, I. N.,1999. A semiparametric two-step estimator in a multivariate long memory model. Journal of Econometrics 90, 29-153.

Lobato, I. N., Robinson, P. M., 1996. Averaged periodogram estimation of long memory. Journal of Econometrics 73, 303-324.

Lobato, I. N., Robinson, P., 1998. A nonparametric test for I(0). Review of Economic Studies 68, 475-495.

Maddison, A., 1995. Monitoring the world economy 1820-1922. Development Centre of the Organization for Economic Cooperation and Development, Paris.

Maddison, A., 2001. The world economy in millienial perspective. Organization for Economic Cooperation and Development, Paris.

Mandelbrot, B. B., 1972. Statistical methodology for non periodic cycles: from the covariance to R/S analysis. Annals of Economic and Social Measurement 1, 259-290.

Mandelbrot, B. B., 1975. A fast fractional gaussian noise generator. Water Resources Research 7, 543-553.

Mandelbrot, B. B., Wallis, J., 1968. N. Joseph and operational hydrology. Water Resources Research 4, 909-918.

Marinucci, D., 2000. Spectral regression for cointegrated time series with long-memory innovations. Journal of Time Series Analysis 21, 685-706.

McLeod, A. I., Hipel, K. W., 1978. Preservation of the rescaled adjusted range, 1. A reassessment of the Hurst phenomenon. Water Resources Research 14, 159-178.

Mikosch, T., Starica, C., 2004. NOnstationarities in financial time series, the long range dependence and the IGARCH effects. Review of Economics and Statistics 86, 378-390.

Montanes, A., Olloqui, I., Calvo, E. 2004. Selection of the break in Perron-type tests, this issue of the Journal.

Nelson, C. R., Plosser, C.I., 1982. Trends and random walks in macroeconomic time series. Journal of Monetary Economics 10, 139-162.

Nielsen, M. O., 2004. Optimal residual based tests for fractional cointegration and exchange rate dynamics. Journal of Business and Economic Statistics 22, 331-345.

Nyblom, J., 1989. Testing for the constancy of parameters over time. Journal of the American Statistical Association 84, 223-230.

Ohanissian, A., Russell, J. R., Tsay, R. S., 2003. Using temporal disaggregation to distinguish between true and spurious long memory. Mimeo, University of Chicago.

Pai, J., Ravishanker, S., N.,1998. Bayesian analysis of autoregressive fractionally integrated moving average processes. Journal of Time Series Analysis 19, 99-112.

Parke, W. R., 1999. What is fractional integration?. The Review of Economics and Statistics 81, 632-638.

Perron, P., 1989. The great crash, the oil price shock and the unit root hypothesis. Econometrica, 57, 1361-1401.

Perron, P., 1997. Further evidence from breaking trend functions in macroeconomic variables. Journal of Econometrics 80, 355-385.

Perron, P., Vogelsang, T.J. 1992. Testing for a unit root in a time series with a shift in mean, corrections and extensions. Journal of Business and Economic Statistics 10, 467-470.

Perron, P., Zhu, X., 2004. Structural breaks with deterministic and stochastic trends. this issue of the Journal.

Peters, A., Sibbertsen, P. 2001. Robust tests on fractional cointegration. University of Dortmund SFB 475 Technical Report 29/2001.

Pesaran, H., Timmerman, A. 2004. Small sample properties of forecasts from autoregressive models under structural breaks, this issue of Journal.

Phillips, P. C. B., Durlauf, S.J., 1986. Multiple time series regression with integrated processes. Review of Economic Studies 53, 473-495.

Phillips, P. C. B., Shimotsu, K., 2004a. Local Whittle estimation in nonstationary and unit root cases. Annals of Statistics 32, 656-692.

Phillips, P. C. B., Shimotsu, K., 2004a. Exact local Whittle estimation of fractional integration. Cowles Foundation Discissuion Paper N. 1367 (Revised).

Quandt, R.E., 1960. Tests of the hypothesis that a linear regression obeys two separate regimes. Journal of American Statistical Association 55, 320-330.

Rappoport, P., Reichlin, L., 1989. Segmented trends and nonstationary time series. The Economic Journal 99 (Conference Volume), 168-177.

Resnick, 1987. Extreme values, regular variation and point processes. Springer-Verlag, New York, NY.

Robinson, P. M., 1978. Statistical inference for a random coefficient autoregressive model. Scandinavian Journal of Statistics 5,163-168.

Robinson, P. M., 1991. Testing for strong serial correlation and dynamic conditional heteroscedasticity in multiple regression. Journal of Econometrics 47, 67-84.

Robinson, P. M., 1994a. Time series with strong dependence, in (ed) Sims, C. A., Advances in Econometrics', Sixth World Congress, Vol. 1. Cambridge University Press, Cambridge.

Robinson, P. M., 1994b. Semiparametric analysis of long-memory time series. The Annals of Statistics 22, 515-539.

Robinson, P. M., 1994c. Efficient tests of nonstationary hypotheses. Journal of the American Statistical Association 89, 1420-37.

Robinson, P. M., 1995a. Log-periodogram regression of time series with long range dependence. The Annals of Statistics 23, 1048-1072

Robinson, P. M., 1995b. Gaussian semiparametric estimation of long range dependence. The Annals of Statistics 23, 1630-1661.

Robinson, P. M., 2003. Long memory time series, in (ed) Robinson, P.M., Time Series with Long Memory, Oxford University Press, Oxford.

Robinson, P. M., Hualde, J. 2003. Cointegration in fractional systems with unknown integration orders. Econometrica 71, 1727, 1766.

Robinson, P. M., Iacone, F., 2004. Cointegration in fractional systems with deterministic trends, this issue of the Journal.

Robinson, P.M., Marinucci, D., 2000. The averaged periodogram for nonstationary vector time series. Statistical Inference for Stochastic Processes 3, 149-160.

Robinson, P.M., Marinucci, D., 2001. Narrow-band analysis of nonstationary processes. The Annals of Statistics 29, 947-986.

Robinson, P. M., Y. Yajima., 2002. Determination of cointegrating rank in fractional systems. Journal of Econometrics 106, 217-241.

Robinson, P. M., Zaffaroni, P. ,1997. Modelling nonlinearity and long memory in time series. Fields Institute Communications 11, 161-170.

Robinson, P. M., Zaffaroni, P., 1998. Nonlinear time series with long memory: a model for stochastic volatility. Journal of Statistical Planning and Inference 68, 359-371.

Rosenblatt, M., 1956. A central limit theorem and a strong mixing condition. Proceedings of the National Academy of Sciences 42, 43-47

Samuelson, P. A., 1973. Proof that properly discounted present values of assets vibrate randomly. The Bell Journal of Economics and Management Science 4, 369-374.

Sowell, F. B., 1990. The fractional unit root distribution. Econometrica 58, 495-505.

Sowell, F. B., 1992. Maximum likelihood estimation of stationary univariate fractionally integrated time series models. Journal of Econometrics 53, 165-188.

Stock, J.H., Watson, M.W., 1988. Testing for common trends. Journal of the American Statistical Association (Theory and Methods) 83, 1097-1107.

Stock, J.H., Watson, M.W., 2002. Macroeconomic forecasting using diffusion indexes. Journal of Business and Economic Statistics 20, 147-162.

Szego, G., 1975. Orthogonal polynomials. American Mathematical Society, 4th Edition.

Taqqu, M. S., 1975. Weak convergence to fractional brownian motion and to the Rosenblatt process. Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete 31, 287-302.

Taqqu, M. S., Teverovsky, V., Willinger, W., 1995. Estimators for long range dependence: an empirical study. Fractals 3, 785-798.

Tsay, W-J., 2000. Long memory story of the real interest rate. Economics Letters 67, 325-330.

Velasco, C., 1999. Gaussian semiparametric estimation of non-stationary time series. Journal of Time Series Analysis 20, 87-127.

Velasco, C., 2003. Gaussian semi-parametric estimation of fractional cointegration. Journal of Time Series Analysis 24, 345:378.

Velasco, C., Robinson, P. M., 2000. Whittle pseudo-maximum likelihood estimation for nonstationary time series. Journal of the American Statistical Association 95, 1229-1243.

Vogelsang, T.J., Perron, P., 1998. Additional tests for unit root allowing the possibility of breaks in the trend function. International Economic Review 39, 1073-1110.

Whitcher, B., 2004. Wavelet-based estimation for seasonal long-memory processes. Technometrics 46, 225-238.

Whittle, P., 1951. Hypothesis testing in time series analysis. Almquist and Wiksells, Uppsala.

Wright, J. H., 1998. Testing for a structural break at unknown date with long-memory disturbances. Journal of Time Series Analysis 19, 369-376.

Wright, J. H., 2000. Log-periodogram estimation of long memory volatility dependencies with conditionally heavy tailed returns. FRB International Finance Discussion Paper No. 2000-685.

Young, C.H., 1974. Asymptotic behaviour of trigonometric series. Chinese Univ. Hong Kong.

Zaffaroni, P., 2003. Gaussian inference on certain long-range dependent volatility models. Journal of Econometrics 115, 199-258.

Zivot, E., Andrews, D. W. K., 1992. Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. Journal of Business and Economic Statistics 10, 251-270.