Columbia University

Department of Economics
Discussion Paper Series

The Long Range Dependence Paradigm for Macroeconomics and Finance

Marc Henry
Paolo Zaffaroni

Discussion Paper #:0102-19

Department of Economics
Columbia University
New York, NY 10027

March 2002

The Long Range Dependence Paradigm for Macroeconomics and Finance

Marc Henry, Paolo Zaffaroni

ABSTRACT The long range dependence paradigm appears to be a suitable description of the data generating process for many observed economic time series. This is mainly due to the fact that it naturally characterizes time series displaying a high degree of persistence, in the form of a long lasting effect of unanticipated shocks, yet exhibiting mean reversion. Whereas linear long range dependent time series models have been extensively used in macroeconomics, empirical evidence from financial time series prompted the development of nonlinear long range dependent time series models, in particular models of changing volatility. We discuss empirical evidence of long range dependence as well as the theoretical issues, both for economics and econometrics, such evidence has stimulated.
Keywords and phrases. Self-similarity, time series, FARIMA, nonlinear time series, ARCH, stochastic volatility, arbitrage.

1 Introduction

[48] first pointed out that nonparametrically estimated power spectra of many economic variables, such as industrial production and commodity price indexes, suggested overwhelming importance of the low frequency components. [84] observed a self-similar behaviour in the distribution of speculative prices, and proposed continuous and discrete time fractional models, such as fractional Brownian motion [85]. He then developed the Hurst rescaled range (hereafter R/S) analysis, originally introduced in [83], for social measurement purposes. However, initial sizeable empirical success of the long-range dependence (LRD) concept in economics is certainly related to the autoregressive fractionally integrated moving average model (hereafter ARFIMA), proposed simultaneously by [64] and [51], and the increased flexibility it brings to the Box-Jenkins linear time series methodology, in particular for modeling macroeconomic time series since [30]. The relative success of the LRD concept in economics may also be attributed to the development of a rationale for its presence in macro-level economic and financial systems based on the aggregation of micro units. This was originally proposed in [95] and further developed in [49], both in the context of contemporaneous aggregation of heterogeneous AR(1) micro-units.

Estimation of fully parametric long memory models such as ARFIMA appeared cumbersome, especially in the time domain, and, moreover, practical estimation

AMS Subject classification. Primary 62M10; secondary 60G18,91B28.

of these models indicated that the estimated values of the long memory parameter was largely affected by the parametric specification of the short run dynamics part of the model, i.e. the ARMA component. Hence, an important reason which gave further impetus to the importance of LRD in economics, was the introduction by [40] of a semiparametric estimator (hereafter GPH) for the estimation of the degree of LRD in a time series. In fact, GPH is robust to many forms of complicated short run dynamics insofar as it is based on a frequency domain ordinary least squares regression of periodogram ordinates in a shrinking band of frequencies around zero. Although not formally developed, this possibility was first suggested by [51] and [68]. Due to its simplicity of implementation (it is based on a simple univariate linear regression), this estimation method was extensively applied to macroeconomic and financial time series well before of any satisfactory theoretical analysis of its asymptotic distributional properties, finally provided by [102]. More efficient and robust semiparametric estimators of long memory have then been proposed by [69], [99], [62], [65], [89], [4], [105] and others.

Applications to macroeconomic time series, discussed in section 2, are intimately linked with the extensive unit root literature and the use of a fractional specification within the linear Box-Jenkins representation to account for high persistence of shocks coupled with possible mean reversion of the levels (i.e. the observations themselves) in series of income, consumption and prices (inflation and exchange rates).

Empirical finance research had a tremendous impact in emphasizing the importance of the LRD paradigm, both empirically and theoretically. On the one hand, the availability of very large time series of high-frequency data, allowed easier and more convincing detection of long memory. On the other hand, some empirical findings prompted the development of nonlinear time series models apt to fit the empirical distribution of asset returns, synthesized in a number of well known stylized facts, including slow decay of sample autocorrelations of squared returns. This has stimulated research aiming at establishing asymptotic theory for such models and deriving their implications in terms of asset pricing and risk management in arbitrage-free theoretical frameworks. These issues are the focus of section 3.

2 Applications of Long Range Dependence to Macroeconomic Time Series

2.1 The Fractional Model

For the sake of discussing macroeconomic applications, the LRD paradigm may be suitably summarized with the following nonparametric specification for the time series Y_t, $t = 0, \pm 1, \ldots$, where the usual notation for income is used:

$$\begin{aligned} Y_t &= \mu + (1-L)^{-d} X_t \\ X_t &= \sum_{j=0}^{\infty} \alpha_j \varepsilon_{t-j} \end{aligned} \tag{2.1}$$

where $\alpha_0 = 1$ and $\sum_{j=0}^{\infty} |\alpha_j| < \infty$, and the ε_j's follow a martingale difference sequence. This specification includes the ARFIMA model in the case where

$$\alpha(z) = \sum_{j=0}^{\infty} \alpha_j z^j = \frac{b(z)}{a(z)} \tag{2.2}$$

where $a(z)$ and $b(z)$ are finite order polynomials with zeros outside the unit circle in the complex plane. Other parametric specifications of LRD include fractional Gaussian noise and an extension of the Bloomfield exponential model as in 3.19 of [100]. Neither, however, has shared the empirical success of the ARFIMA model.

The process Y_t is denoted $\mathrm{I}(d)$ in reference to the degree of integration, and fundamental properties of the series Y_t can be described in terms of interval regions for the parameter d. Y_t is covariance stationary for $d < 1/2$ and invertible for $d > -1/2$. In such cases, let $f(\lambda)$ be the spectral density of the process satisfying

$$\rho_k := \mathrm{cov}(Y_t, Y_{t+k}) = \int_{-\pi}^{\pi} f(\lambda)\cos(k\lambda)d\lambda \tag{2.3}$$

for $k = 0, \pm 1, \ldots$. When $d = 0$, the spectral density is finite and positive at zero and autocorrelations are summable. Usual rules of inference apply, in particular relative to moment estimation. The case where $d < 0$, called antipersistence, is characterized by a shrinking spectral density at zero frequency, and it is empirically relevant to the extent that it describes the behaviour of overdifferenced series, i.e. the result of first differencing in series that were mistakenly believed to have a unit root (i.e. a degree of integration $d = 1$). LRD proper arises when $d > 0$ and it is characterized by a singularity in the spectrum at frequency zero

$$f(\lambda) \sim c\lambda^{-2d} \text{ as } \lambda \to 0^+ \tag{2.4}$$

and, correspondingly, nonsummable autocovariances following

$$\rho_k \sim ck^{2d-1} \text{ as } k \to \infty. \tag{2.5}$$

(2.4) and (2.5) can be used as alternative (though not equivalent, see [122]) definitions of LRD. When $d \geq 1/2$, the variance of the process is infinite, and a non-integrable pseudo-spectrum can be defined with power also concentrated at the origin; and the unit root case is embedded as a special case of specification 2.1. A particularly interesting region for macroeconomic applications is $1/2 < d < 1$ where the time series Y_t has infinite variance, it displays a high degree of persistence as exemplified by 2.5 and yet mean reverts in the sense, for instance, that the impulse response function is slowly decaying. In particular, shocks do not have full persistence because for $d < 1$, $(1-z)^{1-d}\alpha(z)$ has a unit root, so that the long run impulse response of the system is zero.

2.2 Estimation of Long Range Dependence in Linear Models

Consider a sample of size n from the process Y_t. A survey on asymptotic theory relevant (in particular) to parametric estimation of the ARFIMA model is given in [100]. Exact maximum likelihood is efficient in the Fréchet-Darmois-Cramér-Rao (hereafter FDCR) sense when the ε_t's are normally and identically distributed ([26]). The Whittle approximate likelihood is asymptotically efficient under Gaussianity ([38]) and remains $\sqrt{n}$-consistent and asymptotically normal for possibly non-normal identically and independently distributed ε_t's ([46]). A result that is particularly relevant to macroeconomic time series is the apparently greater robustness properties of the Whittle estimate in small samples when the mean is unknown ([21]).

Parametric exact and approximate maximum likelihood estimation, however, relies heavily on a correct specification of the short run dynamics, i.e. the $a(z)$ and $b(z)$ polynomials in 2.2. The semiparametric approach advocated by [51] and [68], and developed by [40], [69], [99], [102], [101] and others, relies only on (2.1) or, more generally, on (2.4), which only specifies the spectral density in a neighbourhood of zero, the frequency of interest. Like the Whittle likelihood, they are periodogram based, but, in accordance with the local specification, only a small proportion of all periodogram ordinates are used in the estimation of d. The choice of bandwidth m, the number of periodogram ordinates $I(\lambda_j) = (1/2\pi n)\,|\sum_{t=1}^{n} Y_t \exp(it\lambda_j)|^2$ used in the estimation, with $\lambda_j = 2\pi j/n$ and $1/m + m/n \to 0$, is crucial, as it drives the efficiency of the $\sqrt{m}$-consistent estimates. See [58] for a discussion of optimal and feasible bandwidth choice in the stationary region for those of the above mentioned estimates that are $\sqrt{m}$-consistent and asymptotically normal.

Even though the local Whittle proposed by [69] is more efficient (see [101] for asymptotic theory and [77] for a multivariate extension), and more robust (see [104]), only the GPH has been extensively used by practitioners*. [30] and [31] apply it to the analysis of output and consumption, [20] to exchange rates, [56] to inflation rates. In the first two cases, $\sqrt{n}$ order bandwidths are used. This ad hoc bandwidth choice was suggested by [40] for the stationary region. [66] prove that the mean squared error minimizing bandwidth m is of order $n^{4/5}$ which is the rate upper bound for its class of estimators as shown in [44]. However, this concerns the stationary region $d < 1/2$, and most macroeconomic time series seem to be nonstationary.† [56] report results with choices of m of order of magnitude n^{α}, $1/2 < \alpha < 1$, and they find little variation in the estimated parameter values.

As mentioned earlier, the parameter region $1/2 < d < 1$ where the process is infinite variance but mean reverting is of particular interest for the analysis of macroeconomic time series. [118] extended the GPH asymptotic normality result of [102] to that region for Gaussian series. However, the behaviour of the GPH across the unit root $d = 1$ value is problematic, as Monte Carlo results from [67] point to a possible inconsistency of the GPH for $d > 1$. [119] proposes a modified

*[105] propose a unifying framework for the GPH and the local Whittle estimates within a class of M-estimates, and improve efficiency with the use of higher-order kernels.

†In any case, orders of magnitude are poor guides in small samples when the constant is unknown. Indeed, [113] argues that the [30] result showing strong mean reversion in output is due to a mispecification of short run dynamics through too large a choice of bandwidth ($m = 11$ for a sample size of $n = 272$).

asymptotically normal version based on smoothing and tapering which robustifies the estimate against linear trends -a particularly appealing feature given that one of the major applications of the LRD paradigm has been to contribute to the debate as to whether GNP is difference stationary or trend stationary. As for small sample behaviour of the GPH and other semiparametric estimates as compared to fully parametric ones, the crucial trade-off between robustness and efficiency in the choice of bandwidth[‡] had been little documented in small sample Monte Carlo investigations (see for instance [114], [58]).

2.3 Long Range Dependence in Aggregate Macroeconomic Series: Evidence and Rationale

One of the major debates in macroeconomic research concerns the persistence of economic shocks on income and the nature of the long cycles observed in US output (Kondratiev, Kuznets or Juglar cycles) and documented in [70] and others. [1] was the first to propose the use of the LRD paradigm to analyze such long cycles and other low frequency characteristics of income series. Now, as [48] noted, one of the more pervasive characteristics of macroeconomic time series is a concentration of power (or variance) in low frequencies, which he dubbs "the typical spectral shape of economic variables". As [100] notes, such explosion of spectral estimates at zero frequency would be consistent with the presence of a unit root (and total persistence of shocks) in the levels of the series, if it were not often associated, as is the case with the Beveridge Wheat Price Index ([11]), with first differenced series with very low power around zero frequency, consistent with I(d) behaviour, $d < 0$. This would suggest that the original series were integrated of order less than one. Besides, unlike the case of financial series with efficient markets hypotheses, there are no theoretical underpinnings for an exact unit root in macroeconomic time series as opposed to mean reverting fractional alternatives. I(d) behaviour with $0 < d < 1$ is also suggested by the conflicting findings concerning aggregate price series from nine industrialized countries analyzed in [8], where the null of a unit root is rejected by Augmented Dickey-Fuller procedures and the null of I(0) is rejected by KPSS tests.

An immediate explanation for evidence of LRD in aggregate price series is the inheritance through agricultural prices of the dependence structure of geophysical time series such as rainfall, riverflow and climactic series in which [83], [72], [63] among others have found LRD to be pervasive. LRD shocks (possibly inherited from underlying geophysical processes) in a real business cycle model of the economy (see [71]) can account in the same way for the presence of LRD in aggregate income series. However, a more appealing explanation for long memory in income series is based on aggregation of output from a large number of sectors each submitted to white noise shocks. [5] similarly justify the presence of LRD in inflation series.[§]

[‡]Reducing bandwidth increases the robustness of asymptotic results not only against misspecification of short run dynamics, but also against departures from Gaussianity in the ε_t's (see [104] and [19]).

[§]In most existing literature on interest rates, the short rate process has the property

These justifications are based on an aggregation result in [95] and [49] with the more general implication that LRD may be found in the aggregate produced from a large number of heterogeneous autoregressive processes describing the microeconomic dynamics of each unit, e.g. the behaviour of each agent when heterogeneity among agents is allowed. In the generalized framework of [75], the results involving the memory can be summarized as follows. Focusing on AR(1) for simplicity's sake, let $Y_{n,t} = (1/n)\sum_{i=1}^{n} y_{i,t}$ be the result of the aggregation of the n micro variables

$$y_{i,t} = \alpha_i y_{i,t-1} + u_t + \varepsilon_{i,t} \quad i = 1, \ldots, n \tag{2.6}$$

where u_t is a common shock and $\varepsilon_{i,t}$ is an idiosyncratic shock. The latter represents a source of time-varying heterogeneity. Assume that the shocks are i.i.d. white noise mutually independent. Heterogeneity across the coefficients α_i, assumed i.i.d. drawn from a nonnegative r.v., is defined by means of a mild semiparametric specification of the latter's probability density function, say $B(\alpha)$. As $n \to \infty$, $Y_{n,t}$ displays short range dependence, with exponentially decaying autocorrelation function (ACF), when the support of $B(\alpha)$ is restricted to $[0, \gamma)$, $\gamma < 1$. It may exhibit LRD when $B(\alpha)$ has support $[0, 1)$ and in particular, $Y_{n,t}$ has, asymptotically, infinite variance LRD when $B(\alpha)$ is concentrated enough in the neighbourhood of 1. Within an economic model with heterogeneous agents behaving according to the standard profit-maximization principle, [87] considers an economic model where agents face heterogeneous and time-varying costs of adjustment in setting their optimal level of output. In this context, he finds that cross-sectional distributions with the required shape in order to induce long memory in the aggregate might arise endogenously.

2.4 Long Range Dependence and Persistence

In the vast literature that followed [90], some consensus seems to have arisen on the fact that US postwar GNP is well represented by low order ARIMAs with a single unit root, mostly as a result of failure to reject null I(1) hypotheses against autoregressive alternatives . However, the non standard asymptotics and lack of Pitman efficiency (as noted in [42]) as well as evidence in [31] that some of the traditional methods have low power against fractional alternatives, have prompted the use of procedures based on LRD specifications with continuous parameter values across the value of interest $d = 1$, such as [98]. [30], [113] and [42] typically fail to discriminate between a unit root and mean reverting

that correlation between short rates n periods apart goes to zero exponentially with n. This is not supported by the data, as considerable residual variation is found in long yields. [5] partially resolve the puzzle with a long memory modeling of the short rate process and justify this assumption as either inherited from long memory in money growth, inflation and monetary policy in general, or as the effect of aggregation across agents with heterogeneous beliefs. They use a fully specified time domain ML procedure and cite [112].

See [93] for a survey on unit root testing fractional alternatives in a variety of US income series (including an extended version of the [90] data) due to the lack of data points available.

In spite of inconclusive inference results, the introduction of the LRD paradigm in the debate over the persistence of output has been instrumental in partially resolving empirical inconsistencies resulting from the unit root (and fully persistent) representation for output. First of all, conflicting evidence on persistence measures from unrestricted ARIMA and Unobserved Component models (in [18] and [120]) is discussed in a LRD framework by [30]. The second empirical inconsistency we mention a little more at length: it concerns the conflicting evidence of a unit root in income and the convergence of output to a steady state. Most empirical studies based on growth regressions point to a 2% uniform exponential rate of convergence to a long run steady state (see for instance [9] and [86]). This finding is inconsistent with full persistence of shocks implied by a unit root in income, but not with the type of persistence implied by infinite variance but mean reverting LRD income, as demonstrated by [88] in the framework of a Solow-Swan growth model. They show that an I(d) income with $1/2 < d < 1$ could produce spurious evidence of uniform exponential convergence (as in the data considered by previous studies), while the actual hyperbolic rate of convergence would be too slow for standard unit root tests to reject the null of a unit root against covariance stationary alternatives.[‖]

The third empirical inconsistency we mention is the excessive smoothness observed in consumption series compared to the implications of the Permanent Income Hypothesis (hereafter PIH) based on a unit root specification for income and under rational expectations. This excess smoothness, or Deaton's Paradox, is presented in [27] who stresses that, with a fully persistent specification of income, the PIH under rational expectations implies that innovations to consumption should have larger variance than innovations to income. More precisely, if income follows model 2.1 with $d = 1$, then the PIH predicts (see for instance [53]) that consumption will react to innovations on income as $\Delta C_t = \kappa_\infty \varepsilon_t$, where $\kappa_\infty = \sum_{j=0}^{\infty} \beta^j \alpha_j$ is the long run discounted impulse response (β is the discount factor). [27], and [17] show that under a variety of ARIMA specifications for income and reasonable assumptions on the interest rate, κ_∞ is generally larger than 1, which is another way of stating Deaton's Paradox. [32] point out that if $d < 1$, however, the impulse response function is decaying, which allows for a smoothed response of consumption to income shocks, which is more in line with the rationale of the PIH. [57] presents a similar resolution of the Deaton Paradox with postulated values of d, and tests whether these paradox resolving values of d are consistent with observed income processes. However, he uses a test based on the [76] modified R/S statistic which has no clear optimality properties (see [100]), and seems to have undesirable small sample properties (see [117]).

[‖]Such a specification for income also accounts for the rejection of convergence in [10] based on cointegration tests, as, in case income is I(d), $d < 1$, the latter may be misspecified.

2.5 Long Range Dependence and Co-persistence

Within the LRD paradigm, the standard definition of cointegration can be extended to allow for fractional integration in both the original variables and the cointegrated residuals, as proposed by [50], [99], [47]. This provides a generalized framework for the analysis of co-persistence which has been particularly fruitful in the description of exchange rates from the post Bretton Woods currency float, and in an empirical reappraisal of the purchasing power parity equilibrium (hereafter PPP). [29] challenge the view that deviations from purchasing power parity can be suitably modeled by martingale differences. They model exchange rates as ARFIMA series which they estimate with the Whittle quasi-likelihood, and find evidence of mean reversion in deviations from parity. A similar study is found in [22]**. [106] discuss parametric and semiparametric procedures for the estimation of various forms of fractional cointegration and establish their asymptotic distributional properties.

3 Long memory implications in finance

3.1 Long memory in asset prices

Following the mixed empirical success of the LRD paradigm in macroeconomics, a great deal of research has focused on possible evidence of LRD in financial asset returns defined as $x_t = \log(P_t/P_{t-1})$, where P_t is the price of the asset. [84] first suggested the possibility that asset prices could exhibit LRD. Using the classical R/S analysis, [52] uncovered significant statistical evidence of LRD in daily returns on securities listed at the New York Stock Exchange. These results were challenged by [76] who proposed a modified R/S statistic, robustifying for possible additional short memory dynamics by taking into account the first q lags of the autocovariance of the observables. He concluded that there is no such clear evidence of LRD in the levels of asset returns. [117], however, show that Lo's modified procedure has dramatic power problems, depending on the assumed degree of short memory in the data. In particular, the test has rapidly decreasing power against LRD alternatives as q increases. [73] use more efficient semiparametric techniques supporting the possibility of LRD, but the results are critically re-examined in [78].

Hence, the empirical evidence concerning LRD in the levels of financial asset returns is far from being clear-cut. Note that most of modern finance theory is based on the martingale difference assumption for the levels of asset returns, in turn a by-product of a martingale assumption for log-prices and therefore the simple presence of autocorrelated returns data in levels, however weak, poses considerable problems. Indeed, there is some empirical evidence of significant autocorrelation at the first few lags often imputed to non-synchronous trading and other market microstructure effects; see e.g. [3] for forex returns and [34] for returns on the Standard & Poor 500 index ($S\&P\,500$).

**See also [23], [6], [20].

There is general agreement on the fact that asset returns exhibit more memory in the squares than in the levels. This stylized fact was refined to the presence of LRD in many nonlinear tranformations of absolute returns (including squared and log-squared returns). Using daily data of the $S\&P\,500$ closing price index, [34] show that there is significant autocorrelation for lags up to 10 years (approximately $2,500$ lags with daily data) when considering the squares return x_t^2 and more generally power transformations such as $|\,x_t\,|^{\alpha}$ for various positive values of α. Considering a suitable $\alpha < 2$, the ACF of $|\,x_t\,|^{\alpha}$ would be well defined even when the x_t have unbounded fourth moment. A slow rate of decay of shocks to the conditional variance of the $S\&P\,500$ is also found in [39]. This pattern is consistent with theoretical autocovariances of the x_t^2 decaying hyperbolically.

This finding has been corroborated since the availability of high frequency tick-by-tick data. [59], [60], [3] and [2] consider the returns from using the foreign exchange spot rates DM/$ and Yen/$ and, using semiparametric estimators such as the GPH, the local Whittle and the average periodogram estimator of [99], find evidence of covariance stationary LRD in the squared, log-squared and absolute returns. [59] and [60] use the optimal bandwdith selection procedures discussed in [58]

Since [115], it has been thought that the existence of a stable relationship between stock returns and trading volumes could be exploited in forecasting speculative prices dynamics. Using daily data for the thirty stocks which contribute to the Dow Jones industrial average index, [79] provides empirical evidence of long memory in stock return volatility and trading volume, with the same degree of memory, although the hypothesis of fractional cointegration was rejected. A semiparametric estimator of the coherency was used.

Unlike macroeconomic time series, financial data (particularly high-frequency) makes it very natural to use semi and nonparametric estimators, as the usual efficiency loss is compensated by the possibility of using very long spans of data. This, in part, explains why fully parametric long memory models have not been advocated in order to detect, as a preliminary analysis, long memory in asset returns, using in turn the levels and the squares as observables. However, once the general statistical properties of the data are uncovered, precise and efficient estimation of the volatility of asset returns is crucial for valuation of derivative securities linked to such assets (such as options and futures). Furthermore, accurate forecasting of volatility of financial assets is a key ingredient for risk mananagement. Nonparametric probability density estimators for long memory variates (see [96]) do extend to estimation of conditional moments but only for a finite number of lagged dependent variables. On the other hand, semiparametric long memory estimators, such as [102] and [101], focus precisely on long-horizon past information but ignore short run dynamics. Therefore, in order to model the intrinsically non-Markovian dynamics of LRD in conditional variance, and for practical pricing of financial assets, fully parametric long memory models of changing volatility must be considered.

3.2 Long memory volatility models

Financial asset returns display a number of regularities across assets and periods, in particular dynamic conditional heteroskedasticity. This clearly cannot be accounted for by linear models. Furthermore, as originally pointed out by [37] and [12], the data exhibit non-Gaussian kurtosis and dynamic asymmetries such as the 'leverage effect', expressed by a negative cross-correlation between current returns (x_t) and future squared returns (x_{t+u}^2 with $u > 0$). The ARCH model of [35] represents the most famous nonlinear time series model apt to account for such features, except for dynamic asymmetries. Denoting by $\mathcal{F}_t$ the σ-field of events generated by $\{x_s : s \leq t\}$ and setting $\sigma_t^2 = \mathrm{var}(x_t^2 \mid \mathcal{F}_{t-1})$, the ARCH($p$) postulates that

$$\sigma_t^2 = \omega + \alpha_1 x_{t-1}^2 + ... + \alpha_p x_{t-p}^2 \quad a.s., \tag{3.1}$$

where $\omega > 0$, $\alpha_i > 0 (i = 1, .., p)$. Therefore, the conditional variance is a symmetric function of past observations. A plethora of extensions have been developed in order to take other features into account, in particular dynamic asymmetries; see [13] for a recent survey.

[97] generalized (3.1) to

$$\sigma_t^2 = \tau^2 + \sum_{j=1}^{\infty} \psi_j (x_{t-j}^2 - \tau^2) \quad a.s., \ \sum_{i=1}^{\infty} \psi_I^2 < \infty, \tag{3.2}$$

where $\tau^2 > 0$ and $\psi_j \geq 0$. This includes (3.1) when $\psi_j = 0$ for all $j > p$. [97] introduced the ARCH(∞) model (3.2) and

$$\sigma_t^2 = \left(\omega + \sum_{j=1}^{\infty} \phi_j x_{t-j} \right)^2, \quad a.s., \ \sum_{i=1}^{\infty} \phi_i^2 < \infty, \tag{3.3}$$

as alternative hypotheses in deriving score tests of no-ARCH with optimal efficiency against such alternatives. Note that for (3.3) there is no need to impose nonnegativity on ω and on the ϕ_i. Model (3.3), called linear ARCH (LARCH), was further considered in [45] who characterize the low-order statistical properties of the model. In particular, they establish sufficient conditions for LRD in the x_t^2 and weak convergence of their partial sums to fractional Brownian motion $B_H(t)$.

In order to impose martingale difference levels and autocorrelated squares, the data generating process for x_t is routinely defined by

$$x_t = \epsilon_t \sigma_t, \tag{3.4}$$

where the so-called rescaled innovations ϵ_t are i.i.d. with $E(\epsilon_t) = 0$ and $E(\epsilon_t^2) = 1$.

[107] discuss LRD parameterizations for the ψ_j in (3.2). An important parametric version is

$$1 - \sum_{j=1}^{\infty} \psi_j z^j = (1 - z)^d \frac{b(z)}{a(z)}, \tag{3.5}$$

where $d \geq 0$ and $a(z)$ and $b(z)$ are defined as in (2.2). Setting $\psi(z) = 1 - \sum_{j=1}^{\infty} \psi_j z^j$ and $\phi = \tau^2 \psi(1)$, (3.2) can be rewritten as

$$\sigma_t^2 = \phi + (1 - \psi(L)) x_t^2.$$

Note that in this case (3.5) implies $\phi = 0$.

Relaxing the condition $E(\epsilon_t^2) = 1$, [43] show that when

$$(E(\epsilon_t^4))^{1/2} \sum_{i=1}^{\infty} \psi_i < 1, \tag{3.6}$$

there exists a strictly stationary solution to (3.2) with bounded fourth moment. (3.6) allows for hyperbolically decaying ψ_j, although parameterizations such as (3.5) are excluded when $E(\epsilon_t^2) = 1$. Moreover, they show that (3.6) implies summability of the ACF of the x_t^2, ruling out LRD.

The ARCH(∞) (3.2)-(3.4) with the parameterization (3.5) was considered in [7] but letting ϕ be a free parameter. This implies that when both $d, \phi > 0$ (and $E(\epsilon_t^2) = 1$) the x_t^2 are not covariance stationary. They propose Gaussian pseudo-maximum likelihood estimation (PMLE) of the model. However, no formal asymptotic distribution theory yet exists for the PLME nor for exact MLE (making distributional assumptions on the rescaled innovations ϵ_t). This holds irrespective of whether ϕ is considered a free positive parameter or not and for any chosen finite parameterization of the $\psi_j = \psi_j(\theta)$, θ denoting a p-dimensional parameter, including (3.5).

Indeed, even when focusing on exponentially decaying ψ_j, the asymptotic properties of the Gaussian PMLE are known only for the ARCH(p) (see [121]) and the GARCH(1,1) (see [74] and [80]), obtained when $p = 2$ and $\psi_j = \alpha\beta^{j-1}$ ($j \geq 1$) for some nonnegative parameters α, β satisfying the strict stationarity condition, $E \log(\beta + \alpha\epsilon_t^2) < 0$ (see [91]). In a linear long memory semiparametric framework, [104] show that the asymptotic properties of the local Whittle estimator [69] are unchanged when relaxing conditional homoskedasticity of the linear innovations to ARCH(∞) dynamic conditional heteroskedasticity.

The ARCH(∞), and more generally ARCH-type models, can be seen as nonlinear autoregressive models. One can also consider nonlinear moving average (MA) models. In order to take into account the 'leverage effect' and to relax nonnegativity contraints on the conditional variance parameters, [92] introduced the EGARCH, given by (3.4) and

$$\log \sigma_t^2 = \omega + \sum_{k=1}^{\infty} \beta_k g(\epsilon_{t-k}), \quad \sum_{k=1}^{\infty} \beta_k^2 < \infty. \tag{3.7}$$

The scalar function $g(\cdot)$ defines the so-called news impact curve $g(\epsilon_t) = \theta\epsilon_t + \gamma(\mid \epsilon_t \mid -E \mid \epsilon_t \mid)$, for scalar parameters γ, θ. Dynamic asymmetry requires $\theta \neq 0$ and identifiability $\beta_1 = 1$. [92] mentioned the possibility of long memory parameterizations for the β_k, further developed in [14] who proposed estimation of the model by Gaussian PMLE. However the asymptotic properties of the Gaussian PMLE are unknown both for short and long memory parameterization of the EGARCH model.

An alternative nonlinear MA has been proposed by [107]:

$$x_t = \epsilon_t h_t, \tag{3.8}$$

$$h_t = \rho + \sum_{j=1}^{\infty} \alpha_j \epsilon_{t-j}, \ \sum_{j=1}^{\infty} \alpha_j^2 < \infty, \tag{3.9}$$

generalizing the nonlinear MA(1) of [94]. This yields $\sigma_t^2 = \mathrm{var}(\epsilon_t) h_t^2$. The choice of the α_i defines the memory properties of the x_t^2, which are weaker or equal than the ones of the h_t depending on whether ρ is equal to zero or not. Focusing on a wide class of long memory parameterization, including the case where the α_j are the moving average coefficients of the ARFIMA, [124] establishes $n^{1/2}$-consistency (n denotes sample size) and asymptotic normality of the Whittle estimator of the model, when fitting the spectral density of the squares x_t^2.

The nonlinear AR and MA models just described can be expressed as nonlinear transformations of present and lagged values of the ϵ_t. Therefore, these models could be defined as 'one-shock' models. In principle, one can invert the models and express the unobservable innovation, the ϵ_t, as a nonlinear function of present and lagged values of the observable, the x_t, although formal proofs of invertibility may prove very difficult to establish.

Motivated by the fact that they seem to provide a closer approximation to continuous time pricing formulae of modern finance, Stochastic Volatility (SV) models, introduced by [116], replace (3.4) with

$$x_t = \eta_t \sigma_t, \tag{3.10}$$

and set

$$\log \sigma_t^2 = \mu_0 + \sum_{k=1}^{\infty} \mu_i \epsilon_{t-i}, \sum_{k=0}^{\infty} \mu_i^2 < \infty,$$

where $\{\eta_t\}$ represents an i.i.d. sequence with $E(\eta_t) = 0$, $\mathrm{var}(\eta_t) = 1$, different from the ϵ_t. Hence, SV-type models can be defined as 'two-shock' models. In most cases it is assumed that η_t and ϵ_t are independent of one another, although this rules out the possibility to account for dynamic asymmetries. SV challenge ARCH-type models in modeling the empirical distribution of asset returns, see [41] for a complete survey. Although the decoupling of the two shocks makes SV more attractive for moment formulae, the latent nature of the conditional variance complicates the evaluation of the conditional variance through filtering and, more importantly, of the likelihood function itself.

A long memory SV model was introduced by [55], setting

$$\log \sigma_t^2 - \mu_0 = (1-L)^{-d} \epsilon_{t-1}, \ \ 0 < d < 1/2,$$

and proposed estimation by a discrete Whittle estimator, fitting the model spectral density of log x_t^2. [16] show consistency of this estimator and [28] show asymptotic normality of the GPH estimator for this model. [24] discuss statistical properties and temporal aggregation issues relative to a continuous time long memory SV model.

A 'two-shock' nonlinear MA was introduced in [108] which allows time-varying first and second conditional moments. It is defined by

$$x_t = g_t + \eta_t h_t,$$

where the bivariate process g_t, h_t is independent of the η_t, defined in (3.10). The memory properties of the x_t and x_t^2 are characterized for a general specification of the g_t, h_t. The linear specification, with h_t given by (3.9) and $g_t = \mu + \sum_{i=0}^{\infty} \beta_j \epsilon_{t-j}$, $\sum_{i=0}^{\infty} \beta_i^2 < \infty$, seems the most appealing specification, prior to finite parameter modeling for estimation purposes.

Focusing on Gaussian ϵ_t and η_t, [103] shows how all these SV-type models can be nested in a general framework, setting $x_t = f(\delta_t)$ where δ_t is a q-dimensional stationary Gaussian process and $f(\cdot)$ an arbitrary function $f : R^q \to R$ such that $E(f^2(\delta_t)) < \infty$. An asymptotic expansion for the ACF of the x_t so defined is provided, which allows to distinguish the impact of the nonlinear transformation $f(\cdot)$ and of the memory of the δ_t, respectively, on the memory of the x_t. In particular, assuming the δ_t exhibit LRD, primitive conditions on the function $f(\cdot)$ are established such that the x_t would exhibit the same or a weaker degree of LRD, respectively.

Despite the numerous LRD volatility models so far proposed, very little is known in terms of the possible sources of LRD in squared asset returns. Inspired by the corresponding linear time series analysis (cf. section 2.3), the contemporaneous aggregation mechanism seems to provide an important vehicle. [33] and [3] employ the linear aggregation results to a GARCH and SV setting, respectively. In both cases, the aggregate is not truly defined as the cross-sectional arithmetic average. In a strong GARCH setting, [123] establishes the probabilistic properties of the aggregate, defined as the cross-sectional arithmetic average of heterogeneous GARCH. In particular, it is shown that the aggregation mechanism never induces LRD in the conditional variance of the aggregate.

3.3 Pricing implications

The principle of no arbitrage is the corner-stone of all pricing models of modern finance, since the seminal work of [54]. [84] first analyzed the pricing implications of long memory. In a single asset framework, he assumed that the first difference of the fundamental follows a linear process, $P_0(t) - P_0(t-1) = \sum_{i=0}^{\infty} \alpha_i N(t-i)$, for some i.i.d. white noise sequence $\{N(t)\}$ and $\sum_{i=1}^{\infty} \alpha_i^2 < \infty$. Further, it was assumed that arbitrage would induce the speculative price of the asset, $P(t)$, to be a linear function of the fundamental and to display the martingale property, implying $P(t) - P(t-1) = N(t)(\sum_{i=0}^{\infty} \alpha_i)$ under risk neutrality and a zero risk-free interest rate. Hence, LRD of the differenced fundamental, i.e. non-summability of the α_i, directly implies the impossibility of implementing the no arbitrage principle.

Although with ambiguous results (cf. section 3.1), the empirical research that focused on the possibility of LRD in the levels of asset returns prompted the use of fractional Brownian motion as an alternative to the standard Brownian motion model. However, [81] first pointed out that fractional Brownian motion of type $B_H(t)$ $(0 < H < 1, 0 \leq t \leq 1)$ is not a feasible model for asset returns. In fact, $B_H(t)$ is not a semimartingale unless $H = 1/2$ for which $B_{1/2}(t) = B(t)$. For the case $1/2 \leq H < 1$ see also [25] and more recently [111]. [109] extended the result to the case $0 < H < 1/2$.

Without the semimartingale property, the standard approach for pricing contingent claims, based on the equivalent martingale measure, fails. This result stems from the fact that the order-q variation of $B_H(t)$ satisfies (cf. [109, eq. (2.2)]), as $n \to \infty$,

$$\sum_{j=1}^{2^n} | B_H(j2^{-n}) - B_H((j-1)2^{-n}) |^q \to_p \begin{cases} 0, & qH > 1, \\ \infty, & qH < 1, \end{cases}$$

consistent with a semimartingale behaviour for $H = 1/2$ only. [109] shows that the failure of the semimartingale property characterizes all Gaussian processes $\{X(t)\}$ of the form

$$X(t) = \int_{-\infty}^{t} \phi(t-s)dB(s) - \int_{-\infty}^{0} \phi(-s)dB(s), \;\; t \in R^+,$$

with square integrable kernel $\phi(\cdot)$ satisfying $\phi(t) \sim c\, t^{H-1/2}$ as $t \to 0^+$. $B_H(t)$ represents a particular case setting $\phi(t) = t^{H-1/2}$. As a consequence, it follows that choosing $\phi(t)$ such that $\phi(0) = 1$, $\phi'(0) = 0$ and $0 < \phi''(t)t^{5/2-H} = O(1), \;\; t \to \infty$, one can construct Gaussian processes which exhibit the same LRD as $B_H(t)$ while retaining the semimartingale property [109, section 5]. [110] addressed the issue of arbitrage opportunities in a more general framework, with respect to the class of processes with continuous sample paths and bounded q-variation $(1 \leq q < 2)$. [25] and [111] discuss fractional version of the Black-Sholes option pricing formula, based on $B_H(t)$. These developments, which use $B_H(t)$ for describing dynamics of asset prices, could still be useful, though, as the corner-stone assumption of zero transaction costs is likely to fail in practice.

When considering the possibility of LRD in squared returns with martingale difference levels (cf. section 3.2), the above mentioned implications in terms of the possibility of arbitrage do not apply. Therefore the standard risk-neutral valuation methods can be used. [24] discuss option pricing formulae based on a continuous time long memory SV model.

On the other hand, one must be aware of these implications when considering volatility models characterized by a time-varying conditional mean specified, since [36], as a linear function of the square-rooted conditional variance when the latter displays long memory. [82] and [61] proposed continuous time subordinated processes for financial asset returns, apt to display heavy tails and LRD in absolute-valued and squared returns.

Practical use of long memory asset pricing formulae in empirical applications is not so widespread yet. Using daily data of exchange traded long-term equity anticipation securities on the $S\&P$ 500 index, [15] propose an empirical applications of pricing long term options, parameterizing the conditional variance with the so-called fractionally integrated GARCH ((3.7) coupled with $\sum_{k=0}^{\infty} \beta_k L^k = \psi^{-1}(L)$ as from (3.5)).

Acknowledgments

The first author acknowledges support from a "Prix à Savant" of the Université Catholique de Louvain.

References

[1] Adelman, I. (1956): "Long Cycles: Fact or Artifact," *American Economic Review*, 55, 444 463.

[2] Andersen, T., T. Bollerlsev, F. Diebold, and P. Labys (1999): "The distribution of exchange rate volatility," preprint.

[3] Andersen, T., and T. Bollerslev (1997): "Heterogeneous information arrivals and return volatility dynamics: uncovering the long run in high frequency returns," *Journal of Finance*, 52, 975 1006.

[4] Andrews, D., and P. Guggenberger (2000): "A bias-reduced log-periodogram estimator for the long memory parameter," preprint.

[5] Backus, D. K., and S. E. Zin (1993): "Long memory inflation uncertainty: evidence from the term structure of interest rates," *Journal of Money, Credit and Banking*, 25, 681 700.

[6] Baillie, R., and T. Bollerslev (1994): "Cointegration, fractional cointegration and exchange rate dynamics," *Journal of Finance*, 49, 737 745.

[7] Baillie, R., T. Bollerslev, and H. Mikkelsen (1996): "Fractionally integrated generalized autoregressive conditional heteroscedasticity," *Journal of Econometrics*, 74, 3 30.

[8] Baillie, R., C.-F. Chung, and M. A. Tieslau (1996): "Analysing inflation by the fractionally integrated ARFIMA-GARCH model," *Journal of Applied Econometrics*, 11, 23 40.

[9] Barro, R., and X. Sala-i-Martin (1995): *Economic growth.* McGraw Hill: New York.

[10] Bernard, A., and S. Durlauf (1996): "Interpreting tests of the convergence hypothesis," *Journal of Econometrics*, 71, 161 173.

[11] Beveridge (1925): "Weather and harvest cycles," *Economic Journal*, 31, 429 452.

[12] Black, F. (1976): "Studies of stock market volatility changes," in *1976 Proceedings from the American Statistical Association, Business and Economic Statistics Section*, pp. 177 181.

[13] Bollerslev, T., R. Engle, and D. B. Nelson (1994): "ARCH models," in Handbook of Econometrics, Vol. IV, ed. by R. Engle and D. McFadden. Elsevier, ch. 49, pp. 2959-3038.

[14] Bollerslev, T., and H. O. Mikkelsen (1996): "Modeling and pricing long memory in stock market volatility," *Journal of Econometrics*, 73, 151 184.

[15] Bollerslev, T., and H.-O. Mikkelsen (1999): "Long-term equity anticipation securities and stock market volatility dynamics," *Journal of Econometrics*, 92, 75 99.

[16] Breidt, F., N. Crato, and P. de Lima (1998): "The detection and estimation of long memory in stochastic volatility," *Journal of Econometrics*, 83, 325 348.

[17] Campbell, J., and A. Deaton (1989): "Is consumption too smooth?" *Review of Economic Studies*, 56, 357 373.

[18] Campbell, J., and G. Mankiw (1987): "Are output fluctuations transitory?" *Quarterly Journal of Economics*, 102, 857 880.

[19] Chang, S.-K., and P. C. B. Phillips (1999): "Log-periodogram regression: the nonstationary case," mimeo, Yale University.

[20] Cheung, Y.-W. (1993): "Long memory in foreign exchange rates," *Journal of Business and Economic Statistics*, 11, 93 101.

[21] Cheung, Y.-W., and F. X. Diebold (1994): "On maximum likelihood estimation of the differencing parameter of fractionally-integrated noise with unknown mean," *Journal of Econometrics*, 62, 301 316.

[22] Cheung, Y. W., and K. S. Lai (1993a): "Do gold market returns have long memory," *Financial Review*, 28, 181 202.

[23] Cheung, Y.-W., and K. S. Lai (1993b): "A fractional cointegration analysis of purchasing power parity," *Journal of the American Statistical Association*, 11, 103 112.

[24] Comte, F., and E. Renault (1998): "Long memory in continuous time stochastic volatility models," *Mathematical Finance*, 8, 291 323.

[25] Cutland, N. J., P. Kopp, and W. Willinger (1995): "Stock price returns and the Joseph effect: a fractional version of the Black-Scholes model," in *Seminar on stochastic analysis, random fields and applications, Progress in Probability, vol. 36*, ed. by E. Bolthausen, M. Dozzi and F. Russo. Basel, Boston and Berlin: Birkhauser Verlag, pp. 327-352.

[26] Dahlhaus, R. (1989): "Efficient parameter estimation for self-similar processes," *Annals of Statistics*, 17, 1749 1766.

[27] Deaton, A. (1987): "Life Cycle models of consumption: is the evidence consistent with the theory?," in *Advances in Econometrics, Fifth World Congress,* Volume II, ed. by T. Bewley. Cambridge University Press, pp. 121-146.

[28] DEO, R., AND C. HURVICH (1999): "On the log periodogram regression estimator of the Memory parameter in long memory stochastic volatility models," preprint.

[29] DIEBOLD, F. X., S. HUSTED, AND M. RUSH (1991): "Real exchange rates under the Gold Standard," *Journal of Political Economy*, 99, 1252 1271.

[30] DIEBOLD, F. X., AND G. RUDEBUSH (1989): "Long memory and persistence in aggregate output," *Journal of Monetary Economics*, 24, 189 209.

[31] DIEBOLD, F. X., AND G. RUDEBUSH (1991a): "Is consumption too smooth: long memory and the Deaton paradox," *Review of Economics and Statistics*, 73, 1 9.

[32] DIEBOLD, F. X., AND G. RUDEBUSH (1991b): "On the power of Dickey-Fuller tests against fractional alternatives," *Economic Letters*, 35, 155 160.

[33] DING, Z., AND C. W. J. GRANGER (1996): "Modeling volatility persistence of speculative returns: a new approach," *Journal of Econometrics*, 23, 185 215.

[34] DING, Z., C. W. J. GRANGER, AND R. ENGLE (1993): "A long memory property of stock market returns and a new model," *Journal of Empirical Finance*, 1, 83 106.

[35] ENGLE, R. (1982): "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation," *Econometrica*, 50, 987 1007.

[36] ENGLE, R., D. LILIEN, AND R. ROBINS (1987): "Estimating time-varying risk premia in the term structure: the ARCH-M model," *Econometrica*, 55, 391 407.

[37] FAMA, E. (1965): "The behaviour of stock market prices," *Journal of Business*, 38, 34 105.

[38] FOX, R., AND M. S. TAQQU (1986): "Large sample properties of parameter estimates for strongly dependent stationary Gaussian time series," *Annals of Statistics*, 14, 517 132.

[39] GALLANT, A., P. ROSSI, AND G. TAUCHEN (1993): "Nonlinear dynamic structures," *Econometrica*, 61, 871 907.

[40] GEWEKE, J., AND S. PORTER-HUDAK (1983): "The estimation and application of long memory time series models," *Journal of Time Series Analysis*, 4, 221 238.

[41] GHYSELS, E., A. HARVEY, AND E. RENAULT (1996): "Stochastic Volatility," in Handbook of Statistics, Vol. 14. Elsevier, New York.

[42] Gil-Alaña, L., and P. M. Robinson (1997): "Testing of unit root and other nonstationary hypotheses in macroeconomic time series," *Journal of Econometrics*, 80, 241 268.

[43] Giraitis, L., P. Kokoska, and R. Leipus (2000): "Stationary ARCH models: dependence structure and central limit theorem," *Econometric Theory*, 16, 1 22.

[44] Giraitis, L., P. M. Robinson, and A. Samarov (1997): "Rate optimal semiparametric estimation of the memory parameter of the Gaussian time series with long range dependence," *Journal of Time Series Analysis*, 18, 49 60.

[45] Giraitis, L., P. M. Robinson, and D. Surgailis (1998): "A model for long memory conditional heteroscedasticity," preprint.

[46] Giraitis, L., and D. Surgailis (1990): "A central limit theorem for quadratic forms in strongly dependent linear variables and its application to asymptotic normality of Whittle's estimate," *Probability Theory and Related Fields*, 86, 87 104.

[47] Gonzalo, J., and C. Granger (1995): "Estimation of common long memory components in cointegrated systems," *Journal of Business and Economic Statistics*, 13, 27 35.

[48] Granger, C. W. J. (1966): "The typical spectral shape of an economic variable," *Econometrica*, 34, 150 161.

[49] Granger, C. W. J. (1980): "Long memory relationships and the aggregation of dynamic models," *Journal of Econometrics*, 14, 227 238.

[50] Granger, C. W. J. (1987): "Two papers: Generalized Integrated Processes and Generalized Cointegration," Working Paper 87-20, University of California San Diego.

[51] Granger, C. W. J., and R. Joyeux (1980): "An introduction to long memory time series models," *Journal of Time Series Analysis*, 1, 15 29.

[52] Greene, M., and B. Fielitz (1977): "Long term dependence in common stock returns," *Journal of Financial Economics*, 4, 339 349.

[53] Hansen, L. P., and T. J. Sargent (1981): "A note on Wiener-Kolmogorov prediction formulas for rational expectation models," *Economics Letters*, 8, 260 266.

[54] Harrison, J., and D. Kreps (1979): "Martingales and arbitrage in multiperiod securities markets," *Journal of Economic Theory*, 20, 381 408.

[55] Harvey, A. C. (1998): "Long-memory in stochastic volatility," in *Forecasting volatility in financial markets*, ed. by J. Knight and S. Satchell. London: Butterworth-Heineman.

[56] Hassler, U., and J. Wolters (1995): "Long memory in inflation rates: international evidence," *Journal of Business and Economic Statistics*, 13, 37 45.

[57] Haubrich, J. G. (1993): "Consumption and fractional differencing: old and new anomalies," *Review of Economics and Statistics*, 75, 767 772.

[58] Henry, M. (2001): "Robust automatic bandwidth for long memory," *Journal of Time Series Analysis*, forthcoming.

[59] Henry, M., and R. Payne (1997): "An investigation of long range dependence in intra-day foreign exchange volatility," Discussion Paper 264, Financial Markets Group, London School of Economics.

[60] Henry, M., and R. Payne (1997): "Long memory and stationary cointegration in intra-day foreign exchange volatility," Discussion Paper, London School of Economics.

[61] Heyde, C. (1999): "A risky asset model with strong dependence through fractal activity time," preprint.

[62] Hidalgo, J. (1999): "Estimation of the pole of long memory processes," mimeo, London School of Economics.

[63] Hipel, K. W., and A. I. McLeod (1978): "Preservation of the rescaled adjusted range," *Water Resources Research*, 14, 509 518.

[64] Hosking, J. (1981): "Fractional differencing," *Biometrika*, 68, 165 176.

[65] Hurvich, C. (2000): "Fexp estimation of long memory time series," preprint.

[66] Hurvich, C., R. Deo, and J. Brodsky (1998): "The mean squared error of Geweke and Porter-Hudak's estimator of the memory parameter of a long memory time series," *Journal of Time Series Analysis*, 19, 19 46.

[67] Hurvich, C., and B. Ray (1995): "Estimation of the memory parameter for nonstationary and non invertible fractionally integrated processes," *Journal of Time Series Analysis*, 16, 17 41.

[68] Janacek, G. J. (1982): "Determining the degree of differencing for time series via the log spectrum," *Journal of Time Series Analysis*, 3, 177 183.

[69] Künsch, H. (1987): "Statistical aspects of self-similar processes," in *Proceedings of the First World Congress of the Bernouilli Society*, pp. 67 74. VNU Science Press.

[70] Kuznets, S. (1965): *Economic growth and structure*. Norton: New York.

[71] Kydland, F., and E. C. Prescott (1982): "Time to build and aggregate economic fluctuations," *Econometrica*, 50, 1345 1370.

[72] Lawrence, A. J., and N. T. Kottegoda (1977): "Stochastic modeling of riverflow time series," *Journal of the Royal Statistical Association*, 140, 1 47.

[73] Lee, D., and P. M. Robinson (1996): "Semiparametric exploration of long memory in stock prices," *Journal of Statistical Planning and Inference*, 50, 155 174.

[74] Lee, S.-W., and B. E. Hansen (1994): "Asymptotic theory for the GARCH(1,1) quasi-maximum likelihood estimator," *Econometric Theory*, 10, 29 52.

[75] Lippi, M., and P. Zaffaroni (1999): "Contemporaneous aggregation of linear dynamic models in large economies," preprint.

[76] Lo, A. (1991): "Long term memory in stock market prices," *Econometrica*, 59, 1279 1313.

[77] Lobato, I. N. (1999): "A semiparametric two-step estimator in a multivariate long memory model," *Journal of Econometrics*, 90, 129 153.

[78] Lobato, I. N., and N. E. Savin (1997): "Real and spurious long memory in stock market data," *Journal of Business and Economic Statistics*, 16, 261 268.

[79] Lobato, I. N., and C. Velasco (1999): "Long memory in stock market trading volume," preprint.

[80] Lumsdaine, R. (1996): "Consistency and asymptotic normality of the quasi-maximum likelihood estimator in IGARCH(1,1) and covariance stationary GARCH(1,1) models," *Econometrica*, 64, 575 596.

[81] Maheswaran, S., and C. Sims (1993): "Empirical implications of arbitrage-free asset markets," in *Models, Methods and Applications of Econometrics*, Basil Blackwell.

[82] Mandelbrot, B., A. Fisher, and L. Calvet (1997): "A multifractal model of asset returns," Cowles Foundation Discusssion Paper 1164.

[83] Mandelbrot, B., and J. Wallis (1969): "Robustness of the rescaled range R/S in the measurement of non cyclic long run statistical dependence," *Water Resources Research*, 5, 867 988.

[84] Mandelbrot, B. B. (1971): "When can price be arbitraged efficiently? A limit to the validity of the random walk and martingale models," *Review of Economics and Statistics*, 56, 225 236.

[85] Mandelbrot, B. B., and J. W. V. Ness (1968): "Fractional Brownian motions, fractional noise and applications," *SIAM Review*, 10, 422 437.

[86] MANKIW, N., D. ROMER, AND D. WEIL (1992): "A contribution to the empirics of economic growth," *Quarterly Journal of Economics*, 107, 407 437.

[87] MICHELACCI, C. (1999): "Cross-sectional heterogeneity and the persistence of aggregate fluctuations," Working Paper 9906, CEMFI.

[88] MICHELACCI, C., AND P. ZAFFARONI (2000): "(Fractional) Beta Convergence," *Journal of Monetary Economics*, 45, 129 153.

[89] MOULINES, E., AND P. SOULIER (1999): "Broadband log periodogram regression of time series with long range dependence," *Annals of Statistics*, 27, 1415 1439.

[90] NELSON, C., AND C. PLOSSER (1982): "Trends and random walks in macroeconomic time series: some evidence and implications," *Journal of Monetary Economics*, 10, 139 162.

[91] NELSON, D. B. (1990): "Stationarity and persistence in the GARCH(1,1) model," *Econometric Theory*, 6, 318 334.

[92] NELSON, D. B. (1991): "Conditional heteroscedasticity in asset pricing: a new approach," *Econometrica*, 59, 347 370.

[93] PHILLIPS, P. C. B., AND Z. XIAO (1998): "A primer on unit root testing," Cowles Foundation for Research in Economics.

[94] ROBINSON, P. M. (1977): "The estimation of a nonlinear moving average model," *Stochastic Processes and their Applications*, 5, 81 90.

[95] ROBINSON, P. M. (1978): "Statistical Inference for a Random Coefficient Autoregressive Model," *Scandinavian Journal of Statistics*, 5, 163 168.

[96] ROBINSON, P. M. (1991a): "Nonparametric function estimation for long memory time series," in *Semiparametric and nonparametric methods in econometrics and statistics*, ed. by W. Barnett, J. Powell and G. Tauchen, Cambridge University Press, Cambridge, pp. 189-232.

[97] ROBINSON, P. M. (1991b): "Testing for strong serial correlation and dynamic conditional heteroscedasticity in multiple regression," *Journal of Econometrics*, 47, 67 84.

[98] ROBINSON, P. M. (1994a): "Efficient tests of nonstationary hypotheses," *Journal of the American Statistical Association*, 89, 1420 1437.

[99] ROBINSON, P. M. (1994b): "Semiparametric analysis of long memory time series," *Annals of Statistics*, 22, 515 539.

[100] ROBINSON, P. M. (1994c): "Time series with strong dependence," in *Advances in Econometrics*, ed. by C. Sims, vol. 1, pp. 97 107. Cambridge University Press.

[101] Robinson, P. M. (1995a): "Gaussian semiparametric estimation of long range dependence," *Annals of Statistics*, 23, 1630 1661.

[102] Robinson, P. M. (1995b): "Log periodogram regression of time series with long range dependence," *Annals of Statistics*, 23, 1048 1072.

[103] Robinson, P. M. (1999): "The memory of stochastic volatility models," preprint.

[104] Robinson, P. M., and M. Henry (1999): "Long and short memory conditional heteroscedasticity in estimating the memory parameters of levels," *Econometric Theory*, 15, 299 336.

[105] Robinson, P. M., and M. Henry (2001): "Higher-order kernel semiparametric M-estimation of long memory," preprint.

[106] Robinson, P. M., and D. Marinucci (1998): "Semiparametric frequency domain analysis of fractional cointegration," STICERD, Discussion Paper EM/98/348.

[107] Robinson, P. M., and P. Zaffaroni (1997): "Modeling nonlinearity and long memory in time series," *Fields Institute Communications*, 11, 161 170.

[108] Robinson, P. M., and P. Zaffaroni (1998): "Nonlinear time series with long memory: a model for stochastic volatility," *Journal of Statistical Planning and Inference*, 68, 359 371.

[109] Rogers, L. C. G. (1997): "Arbitrage with fractional brownian motions," *Mathematical Finance*, 7, 95 105.

[110] Salopek, D. (1998): "Tolerance to arbitrage," *Stochastic Processes and their Applications*, 76, 217 230.

[111] Shiryaev, A. N. (1998): "On arbitrage and replication for fractal models," preprint.

[112] Sowell, F. (1992a): "Maximum likelihood estimation of stationary univariate fractionally integrated time series models," *Journal of Econometrics*, 53, 165 188.

[113] Sowell, F. (1992b): "Modeling long-run behaviour with the fractional ARIMA model," *Journal of Monetary Economics*, 29, 277 302.

[114] Taqqu, M. S., and V. Teverovsky (1995): "Estimators for long range dependence: an empirical study," *Fractals*, 3, 785 798.

[115] Tauchen, G., and M. Pitts (1983): "The price variability-volume relationship on speculative markets," *Econometrica*, 51, 485 505.

[116] Taylor, S. (1986): *Modeling financial time series*. Chichester: John Wiley.

[117] Teverovski, V., M. S. Taqqu, and W. Willinger (1999): "A critical look at Lo's modified R/S statistic," *Journal of Statistical Planning and Inference*, 80, 211 227.

[118] Velasco, C. (1999a): "Non-stationary log-periodogram regression," *Journal of Econometrics*, 91, 325 371.

[119] Velasco, C. (1999b): "Nonparametric frequency domain analysis of non-stationary multivariate time series," mimeo, Universidad Carlos III.

[120] Watson, M. (1986): "Univariate detrending methods with stochastic trends," *Journal of Monetary Economics*, 18, 49 75.

[121] Weiss, A. A. (1986): "Asymptotic theory for ARCH models: Estimation and testing," *Econometric Theory*, 2, 107 131.

[122] Yong, C. (1974): *Asymptotic Behaviour of Trigonometric Series.* Chinese University of Hong Kong.

[123] Zaffaroni, P. (1999a): "Contemporaneous aggregation of GARCH processes," preprint.

[124] Zaffaroni, P. (1999b): "Gaussian inference on certain long-range dependence volatility models," preprint.

Marc Henry, Department of Economics, Columbia University, 420 W 118th Street, New York, NY10027, U.S.A., e-mail: marc.henry@columbia.edu

Paolo Zaffaroni, Department of Applied Economics, University of Cambridge, Austin Robinson Building, Sidgwick Avenue, Cambridge, CB3 9DE, UK, e-mail: paolo.zaffaroni@econ.cam.ac.uk