Modelling Exchange Rates Volatility with Multivariate Long-Memory ARCH Processes

by **Gilles Teyssière***
The European Commission, Joint Research Centre
& GREQAM, Marseille

GREQAM DT 97B03
First Draft: January 1997
This revised version: September 2000

Abstract

We propose two multivariate long-memory ARCH models: we first consider a long-memory extension of the restricted constant conditional correlations (CCC) model introduced by Bollerslev (1990), and we propose a new unrestricted conditional covariance matrix model which models the conditional covariances as long-memory ARCH processes. We apply these two models to two daily returns on foreign exchanges (FX) rates series, the Pound-US dollar, and the Deutschmark-US dollar. The estimation results for both models show: (*i*) that the unrestricted model outperforms the restricted CCC model, and (*ii*) that all the elements of the conditional covariance matrix share the same degree of long-memory for the period April 1979-January 1997, i.e., after the European Monetary System inception in March 1979. However, this result does not hold for the period September 1971-January 1997, and the floating period before March 1979. Semiparametric methods confirm that the volatilities and co-volatility of the two FX rates share the same long-range component, and that the break in the long-term structure is likely to be caused by the European Monetary System inception.

Keywords: Long-memory processes, conditional heteroskedasticity, multivariate long-memory ARCH models, multivariate FIGARCH models, semiparametric estimators.
JEL classification: C32, G00.

---

*This is a revised version of the **GREQAM DT 97B03** (October 1997), which is available at the following e-mail address: dtgreqam@ehess.cnrs-mrs.fr
I whish to thank for useful comments and discussions Raj Bhansali, Olaf Bunke, Liudas Giraitis, Piotr Kokoszka, Alan Kirman, Remigijus Leipus, participants to the seminars organized by the Department of Mathematics, University of Liverpool, May 1997; the 6^{th} Annual Symposium of the Society for Nonlinear Dynamics and Econometrics, New York, March 1998; the Weierstrass Institute for Applied Analysis and Stochastics, Berlin, April 1998; the ESRC Econometric Study Group Annual Conference, Bristol, July 1998; the 1998 Econometric Society European Meeting in Berlin. The usual caveats apply.
Correspondence: Joint Research Centre of European Commission, ISIS, TP361. 1, Via E. Fermi, I-21020 Ispra (VA), ITALY.
E-Mail: gilles@ehess.cnrs-mrs.fr, gilles.teyssiere@jrc.it
Home Page: http://members.aol.com/gteyssiere

1 Introduction

Most of the returns on speculative assets $R_t = \ln(P_t/P_{t-1})$, where P_t denotes the asset price at time t, are characterized by the absence of a significant correlation between successive observations, which is a consequence of the martingale property of asset prices. However, their variations are clustered, their conditional variance is predictable, and the power transformations $|R_t|^\delta$, where δ is a positive real number, of these series exhibit a significant persistence and dependence between very distant observations called "long-range dependence" or "long-memory".[1]

Several classes of stochastic processes can mimic these properties: the long-memory ARCH processes introduced in Robinson (1991), the long-memory stochastic volatility models proposed by Harvey (1998) and Breidt, Crato and de Lima (1998), the multifractals model of assets returns developed by Mandelbrot, Fisher and Calvet (1997), the multi-factors models of Gallant, Hsu and Tauchen (1998), and the mixture of ARCH and IGARCH processes, mixing a transitory and a permanent component, by Ding and Granger (1996), and Engle and Lee (1999).

We consider here the class of long-memory ARCH models introduced by Robinson (1991) for testing for dynamic conditional heteroskedasticity, and developed later by Granger and Ding (1995) and other authors.[2] All these models were concerned with the univariate modeling of time series. Empirical evidence on futures data,[3] intra-day foreign exchange (FX) rates returns (Henry and Payne (1997)), and stock indexes (Ray and Tsay (2000)), suggest that some time series share the same long-memory component in their conditional variance. Since some time series are likely to be influenced by the same set of events, it is straightforward to consider long-memory in the conditional second moments in a multivariate framework by also modeling the covariation of the series. Furthermore, a common long-range component or 'co-persistence' in a matrix of conditional variances is of interest for long-term forecasts as the relative variations between these covariances are only transitory.

We first consider the extension of the univariate long-memory ARCH models in the restricted multivariate constant conditional correlations (CCC) framework proposed by Bollerslev (1990). As we observed, using semiparametric estimators, that the volatilities and the 'co-volatility' of some series have a common long-memory component, we propose a new multivariate unrestricted framework, which models the covariances as long-memory ARCH processes. We apply these classes to the bivariate modeling of two series of FX rates returns at daily frequency for the periods September 1971-January 1997 and April 1979-January 1997: the Pound-US dollar, and the Deutschmark-US dollar.

This paper is organized as follows. Section **2** discusses the univariate and multivariate long-memory ARCH models, Section **3** presents an application to a bivariate modeling of FX rates returns. Section **4** concludes.

[1] See Beran (1994) and Robinson (1994) for a survey on long-memory processes.

[2] See Robinson and Zaffaroni (1997), Ding and Granger (1996), Giraitis, Kokoszka and Leipus (2000) for rigorous presentations of ARCH(∞) processes.

[3] See a previous version of this work on futures data. As the sample sizes were a bit short for long-memory analysis, and we turn to the analysis of FX data.

2 Multivariate long-memory ARCH models

2.1 The long-memory property of time series

Let $\{y_t\}$ be a stochastic process defined by the following autoregressive (AR) representation

$$(1-\alpha(L))\, y_t = \varepsilon_t \quad \varepsilon_t \text{ white noise} \tag{1}$$

where $\alpha(L)$ is a lag polynomial. We allow for long-memory in $\{y_t\}$ by considering that the order of $\alpha(L)$ is infinite and that its coefficients α_j have the following rate of decay:

$$|\alpha_j| = O\left(j^{-(1+d)}\right) \quad \text{as } j \to \infty \tag{2}$$

where d is a strictly positive real number, $d \in (0, 1/2)$, which controls the hyperbolic rate of decay of the α_j and then represents the long-memory property of the series. Several polynomials satisfy this property:

- The polynomial associated with the fractional d^{th} difference operator, used for defining the fractional Gaussian noise $I(d)$ process, see Mandelbrot and Van Ness (1968), Granger and Joyeux (1980) and Hosking (1981). This polynomial is characterized by the real parameter d. Thus, $1-\alpha(L) = (1-L)^d$, with

$$\alpha_j = \frac{-\Gamma(j-d)}{\Gamma(-d)\Gamma(j+1)}, \quad \text{with} \quad \alpha_j \sim \frac{-1}{\Gamma(-d)} j^{-(1+d)} \quad \text{as } j \to \infty \tag{3}$$

  where $\Gamma(\bullet)$ denotes the Gamma function, d is called the fractional degree of integration of the I(d) process.

- The polynomials associated with the Gegenbauer process, suggested by Hosking (1981) and independently introduced by Andel (1986) and Gray, Zhang and Woodward (1989). These processes are characterized by two parameters d and θ, which are respectively the degree of persistence and the frequency of the singularity of the process. The coefficients of the Gegenbauer polynomials are defined by

$$\begin{aligned} \alpha_j &= \sum_{k=0}^{[j/2]} \frac{(-1)^k \Gamma(-d+j-k)(2\cos(\theta))^{j-2k}}{\Gamma(-d)\Gamma(k+1)\Gamma(j-2k+1)} \\ \text{with} \quad \alpha_j &\sim \frac{\cos\left((j-d)\theta + (d\pi/2)\right)}{\Gamma(-d)\sin^{-d}(\theta)} \left(\frac{2}{j}\right)^{1+d} \quad \text{as } j \to \infty \end{aligned} \tag{4}$$

  where $[\bullet]$ indicates integer part. When $\theta = 0$, these polynomials reduce to the polynomial of the $I(2d)$ process.

- The polynomials defined by the ratio of two Beta functions, introduced by Ding and Granger (1996)

$$\alpha_j = \frac{B(p+j-1, d+1)}{B(p,d)}, \quad \text{with} \quad \alpha_j \sim \frac{d\Gamma(p+d)}{\Gamma(p)} j^{-(1+d)} \quad \text{as } j \to \infty \tag{5}$$

for $p, d > 0$.[4] If $p + d = 1$, these polynomials reduce to the $I(d)$ polynomial. Since these polynomials are characterized by two parameters, they are more flexible than the polynomial of the $I(d)$ process.[5]

The Gegenbauer polynomials capture long-range dependence with persistent component at frequency θ, and are then used for modeling macroeconomic data. The two other polynomials, the coefficients of which are always strictly positive, have been used for modeling the long range dependence in the conditional variance of some financial assets. They also satisfy

$$\sum_{j=1}^{\infty} \alpha_j = 1 \tag{6}$$

2.2 Long-memory in the conditional variance

Define a conditional heteroskedasticity model as

$$R_t = m(R_t) + \varepsilon_t, \quad \varepsilon_t \sim \text{i.i.d.}\, N(0, \sigma_t^2) \tag{7}$$

where $m(R_t)$ denotes the regression function, the conditional variance σ_t^2 depends on the information set I_t consisting of everything dated $t-1$ or earlier. Engle (1982) proposed the ARCH(p) skedastic function:

$$\sigma_t^2 = \omega + \alpha(L)\varepsilon_t^2 \tag{8}$$

where $\alpha(L)$ is a lag polynomial of order p. Robinson (1991) has considered the general case of conditional heteroskedasticity by resorting to long-memory infinite order lag polynomials, and has proposed the following long-memory ARCH:

$$\sigma_t^2 = \sigma^2 + \sum_{j=1}^{\infty} \alpha_j (\varepsilon_{t-j}^2 - \sigma^2) = \sum_{j=1}^{\infty} \alpha_j \varepsilon_{t-j}^2 \tag{9}$$

for some $\sigma^2 > 0$, where the coefficients $\{\alpha_j\}_{j=1}^{\infty}$ of the infinite order lag polynomial $\alpha(L)$ satisfy equations (2) and (6). Thus, equation (9) defines a long-memory process in the variance, in which shocks on the error terms have a persistent effect on the conditional variance. We first consider for $\alpha(L)$ the polynomial defined by the ratio of two Beta functions. Thus, the long-memory ARCH model becomes

$$\sigma_t^2 = \sum_{j=1}^{\infty} \frac{B(p+j-1, d+1)}{B(p,d)} \varepsilon_{t-j}^2 \tag{10}$$

which is a particular case of the long-memory ARCH model derived by Ding and Granger (1996)

$$\sigma_t^2 = (1-\mu)\sigma^2 + \mu\alpha(L)\varepsilon_t^2 \tag{11}$$

[4] If $p = 0$, $\alpha_1 = 1$, $\alpha_j = 0, \forall j > 1$

[5] Ding and Granger (1996) report examples of decay of this polynomial for several values of p and d.

where μ is a parameter $\in [0,1]$. In our case $\mu = 1$.

This long-memory ARCH is interesting: since the coefficients of $\alpha(L)$ are strictly positive, there is no restriction for insuring the conditional variance to be strictly positive, provided that the ratio of the Beta functions exists, i.e., $p > 0$ and $d > 0$, and at least one of the past error terms ε_t is different from zero, which are rather mild constraints. We extend this model by considering parameterized real powers for both σ_t and ε_t. Thus, the long-memory ARCH model (10) becomes:

$$\sigma_t^\delta = \sum_{j=1}^{\infty} \frac{B(p+j-1, d+1)}{B(p,d)} |\varepsilon_{t-j}|^\delta \tag{12}$$

Long-memory in the conditional variance is also modeled by using the ARMA parameterization of some GARCH type processes, and considering fractional roots in the AR component of this ARMA parameterization. We then obtain the class of Fractionally Integrated GARCH processes, which bridges the gap between GARCH and Integrated GARCH processes. A FIGARCH(l,d,q) process is then defined as

$$y_t = \sigma_t \varepsilon_t \qquad \sigma_t^2 = \frac{\omega}{1-\beta(1)} + \left(1 - \frac{(1-\phi(L))\,(1-L)^d}{(1-\beta(L))}\right) \varepsilon_t^2 \tag{13}$$

where $\beta(L)$ and $\phi(L)$ are lag polynomials of respectively finite order p and q, $\beta(L) = \sum_{j=1}^{p} \beta_j L^j$, $\phi(L) = \sum_{j=1}^{m} \phi_j L^j$, the roots of $1-\beta(L)$ and $1-\phi(L)$ being outside the unit circle.[6]

These ARCH(∞) processes have been designed for representing the long-memory properties of the squared and absolute returns, the autocovariance function of which is typical of long-memory processes:

$$\mathrm{Cov}(R_t^2, R_{t+k}^2) \sim C\, k^{2d-1} \quad \text{as} \quad k \to \infty \tag{14}$$

where C is a positive constant. The covariance function is not absolutely summable. However, according to Giraitis, Kokoszka and Leipus (2000), Kazakevičius and Leipus (2000), Kazakevičius, Leipus and Viano (2000), the long-memory ARCH considered here are not stationary. The equations defining the ARCH(∞) process proposed by Ding and Granger (1996) admit a unique covariance stationary solution if $\mu < 1/\sqrt{E(\varepsilon^2)}$. In that case, this ARCH(∞) is short memory since

$$\mathrm{Cov}(R_t^2, R_{t+k}^2) \sim C\, k^{-d-1} \quad \text{as} \quad k \to \infty \tag{15}$$

which is different from equation (14) and is absolutely summable.

Furthermore, the heuristics arguments for the existence of a strictly stationary solution to the equations (13) given in Baillie *et al.* (1996) are questionable: the authors argue that the coefficients of the polynomial $\alpha(L)$ of a FIGARCH process are dominated by those of an IGARCH process. Given that Bougerol and Picard (1992) and Nelson (1990)

[6] The link between this FIGARCH model and the previous long-memory ARCH model, is given in Robinson and Zaffaroni (1997).

have demonstrated that the IGARCH process is strictly stationary and ergodic, they conclude that the FIGARCH process is also strictly stationary and ergodic. This assertion is incorrect since the coefficients of $\alpha(L)$ for an IGARCH process decrease exponentially, while the coefficients of $\alpha(L)$ for a FIGARCH process decrease hyperbolically. Strict conditions for the existence of stationary solutions to ARCH(∞) equations, requiring some constraints on the coefficients $\alpha(L)$, are given in Giraitis, Kokoszka and Leipus (2000), and very recently in Kazakevičius and Leipus (2000), who extended the results of Nelson (1990) and Bougerol and Picard (1992).

The richer parameterization of the FIGARCH allows a greater flexibility than the long-memory ARCH model presented above. The counterpart of this flexibility is a more restrictive set of conditions for insuring the conditional variance to be strictly positive for all t. While the necessary and sufficient conditions for a FIGARCH$(1, d, 0)$ model are not too constraining: $\beta \leq d \leq 1$ and $\omega > 0$, the sufficient conditions for a FIGARCH$(1, d, 1)$ are more restrictive[7]

$$\omega > 0,\ \beta - d \leq \phi \leq (2-d)/3,\ d\,(\phi - (1-d)/2) \leq \beta(\phi - \beta + d) \tag{16}$$

This set of restrictions also applies to others extensions of GARCH-type models to fractional situations. We alternatively can use a log transformation of the conditional variance, like the FIEGARCH model proposed by Bollerslev and Mikkelsen (1996). However, as mentioned by Pagan (1996), the IEGARCH model is not identifiable by Quasi Maximum Likelihood (QML), thus if d tends to 1, the same problem is likely to occur for a FIEGARCH.

2.3 Multivariate long-memory skedastic functions

Let $\boldsymbol{R_t}$ be a n-dimensional vector long-memory ARCH process

$$\boldsymbol{R_t} = \boldsymbol{m}(\boldsymbol{R_t}) + \boldsymbol{\varepsilon_t}, \quad \boldsymbol{\varepsilon_t} \sim \text{i.i.d.}\, \boldsymbol{N}(\boldsymbol{0}, \boldsymbol{\Sigma_t}) \tag{17}$$

where $\boldsymbol{m}(\boldsymbol{R_t})$ denotes the vector regression function, $\boldsymbol{\varepsilon_t}$ is a n-dimensional vector of Gaussian error terms with conditional covariance matrix $\boldsymbol{\Sigma_t}$. We first consider a long-memory extension of the constant conditional correlations model (CCC) proposed by Bollerslev (1990), where the conditional covariance matrix $\boldsymbol{\Sigma_t}$ has as typical element $s_{ij,t}$, with

$$\begin{aligned} s_{ii,t} &= \sigma^2_{ii,t} = \sum_{k=1}^{\infty} \frac{B(p_i + k - 1, d_i + 1)}{B(p_i, d_i)} \varepsilon^2_{i,t-k}, \quad i = 1, \ldots, n \\ s_{ij,t} &= \rho_{ij} \sigma_{ii,t} \sigma_{jj,t}, \quad i, j = 1, \ldots, n \quad i \neq j \end{aligned} \tag{18}$$

where $\rho_{ij} \in (-1, 1)$ denotes the conditional correlation which is supposed to be constant. This parsimonious diagonal specification insures that the conditional covariance matrix is always positive definite, and that the multivariate long-memory ARCH process is stationary if all the univariate processes in the main diagonal are stationary.[8] However, as

[7] See Bollerslev and Mikkelsen (1996).

[8] See Bollerslev (1990), Bera and Higgins (1993), Bollerslev, Engle and Nelson (1994), Engle and Kroner (1995), or Pagan (1996), for surveys on multivariate ARCH processes.

mentioned above, the long-memory ARCH processes might be non-stationary, and in that case the CCC specification is not helpful for this purpose.

This process can be extended by using a parameterization similar to equation (12) as follows:

$$\begin{aligned} s_{ii,t} &= \sigma^2_{ii,t}, \;\; \sigma^\delta_{ii,t} = \sum_{k=1}^{\infty} \frac{B(p_i+k-1, d_i+1)}{B(p_i, d_i)} \left|\varepsilon_{i,t-k}\right|^\delta, \; i=1,\ldots,n \\ s_{ij,t} &= \rho_{ij}\sigma_{ii,t}\sigma_{jj,t}, \quad i,j=1,\ldots,n \quad i\neq j \end{aligned} \tag{19}$$

We also consider the multivariate CCC–FIGARCH process:

$$\begin{aligned} s_{ii,t} &= \sigma^2_{ii,t} = \frac{\omega_i}{1-\beta_i(1)} + \left(1 - \frac{(1-\phi_i(L))(1-L)^d}{1-\beta_i(L)}\right) \varepsilon^2_{i,t}, \; i=1,\ldots,n \\ s_{ij,t} &= \rho_{ij}\sigma_{ii,t}\sigma_{jj,t}, \quad i,j=1,\ldots,n \quad i\neq j \end{aligned} \tag{20}$$

Lastly, we propose two alternative unrestricted parameterizations for the conditional covariance matrix $\boldsymbol{\Sigma_t}$, with typical element $s_{ij,t}$:

$$s_{ij,t} = \sum_{k=1}^{\infty} \frac{B(p_{ij}+k-1, d_{ij}+1)}{B(p_{ij}, d_{ij})} \varepsilon_{i,t-k}\varepsilon_{j,t-k}, \quad i,j=1,\ldots,n \tag{21}$$

and

$$s_{ij,t} = \frac{\omega_{ij}}{1-\beta_{ij}(1)} + \left(1 - \frac{(1-\phi_{ij}(L))(1-L)^{d_{ij}}}{1-\beta_{ij}(L)}\right) \varepsilon_{i,t}\varepsilon_{j,t}, \quad i,j=1,\ldots,n \tag{22}$$

Since there is no analytically tractable set of conditions for insuring $\boldsymbol{\Sigma_t}$ to be positive definite for all t, we implement the positive definiteness constraint in the estimation procedure by imposing this constraint.

This unrestricted long-memory ARCH process is stationary provided that $\boldsymbol{\Sigma_t}$ is a measurable function on the information set I_t, and trace$(\boldsymbol{\Sigma_t}\boldsymbol{\Sigma_t}^\top)$ is finite almost surely. The latter condition holds if the moments are bounded, i.e., E(trace$(\boldsymbol{\Sigma_t}\boldsymbol{\Sigma_t}^\top)^{\boldsymbol{p}}$) is finite for some p.[9] However, given the lack of theoretical results for univariate long-memory ARCH processes, the issue of stationarity for multivariate processes is still open.

Since we are working with data at daily frequency, we can reasonably assume that the error terms are normally distributed.[10] In that case, the log-likelihood function of the multivariate long-memory ARCH model is:

$$\mathcal{L}_T(\zeta) = -\frac{nT}{2}\ln(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\left(\ln|\boldsymbol{\Sigma_t}| + \boldsymbol{\varepsilon_t}^\top \boldsymbol{\Sigma_t}^{-1}\boldsymbol{\varepsilon_t}\right) \tag{23}$$

where ζ and T respectively denote the set of parameters and the sample size. The robust estimators of the variances are given by the heteroskedastic consistent "sandwich" covariance matrix $T^{-1}\mathcal{H}^{-1}(\hat{\zeta})\,\mathcal{I}(\hat{\zeta})\,\mathcal{H}^{-1}(\hat{\zeta})$, where $\mathcal{H}(\hat{\zeta})$ and $\mathcal{I}(\hat{\zeta})$ are respectively the Hessian and the outer-product-of-the-gradient matrices evaluated at the Quasi Maximum Likelihood (QML) estimates $\hat{\zeta}$.

[9] See Bollerslev, Engle and Nelson (1994).

[10] The departure to normality does not affect the QML results with data at daily frequency. However, this assumption is difficult to maintain with very high-frequency data, the leptokurtosis of which can be captured with t-distributed error terms with $t < 7$ or Generalized Exponentially Distributed error terms.

3 Application to the bivariate modeling of exchange rates

We consider several bivariate long-memory ARCH models for two series at daily frequencies: the log of returns on the Pound-US dollar, and the Deutschmark-US dollar exchange rates, defined by $R_t = 100 \ln(P_t/P_{t-1})$, where P_t denotes the exchange rate at time t. We consider here observations from the 24 August 1971, after the collapse of the Bretton-Wood fixed exchange rate system, until January 1997. Several limited variation exchange rates systems have been instituted after this collapse. In particular, during this floating period, some European currencies were governed by several mechanisms limiting their relative variations. The "Snake in the Tunnel" mechanism allowed these currencies to vary by a maximum of ±1.5% against each other, the "Snake", and by a maximum of ±2.5% against the dollar, "the Tunnel". Belgium, France, Italy, Luxembourg, The Netherlands, and West Germany were the first participants. UK joined one month after and withdrew six weeks afterwards. Later, the reference to the dollar was abandoned, and then the system became the "Snake", in which parities were several times realigned. Later, in March 1979, the European Monetary System (EMS) replaced the "Snake". UK joined the EMS in 1990 and left two years later. Two related works lead us to consider the choice of two periods. Kokoszka and Leipus (2000) proposed a CUMSUM based estimator for change-point in ARCH(∞) processes at time t. This estimator is defined by:

$$\hat{t} = \min_t \left\{ |C_t| = \max_{1 \leq j \leq T} |C_j| \right\} \tag{24}$$

where

$$C_t = \frac{t(T-t)}{T^2} \left(t^{-1} \sum_{j=1}^{t} R_j^2 - \frac{1}{T-t} \sum_{j=t+1}^{T} R_j^2 \right) \tag{25}$$

Kokoszka and Leipus (2000) apply this test to the same data set we use in this work. They detected a change in regime in the ARCH structure after the EMS inception in April 1979: t reaches a maximum for $t = 2100$, which is around September 1979, although the pattern of the plot of $|C_t|$ steeply increases before that date. Märdle, Spokoiny and Teyssière (2000), considered also the same data set for the adaptive estimation of a stochastic volatility model, based on the method developed in Spokoiny (1998). This adaptive method consists in finding the largest interval of homogeneity for the parameters of a stochastic volatility (SV) model. They observed a significant change in the estimated parameters of the SV model at the same time point, for both series. This means that the change point may have occurred before as these adaptive methods detect gradual changes with some delay. Bollerslev (1990) mentioned that a break caused by a change in operation procedures of the US Federal Reserve Bank in October 1979 can be also considered.

We then consider two periods: 24 August 1971-21 January 1997 (6630 observations), after the collapse of the Bretton-Wood system, and the EMS period, April 1979- 21 January 1997 (4646 observations).[11]

[11] In an earlier version of the paper, we also consider other periods before the EMS. Estimation results and tests have shown that these sub-periods are not of interest.

3.1 Semiparametric analysis

In a first approach, we estimate nonparametrically the long-memory component in the volatility of the two series. Since the volatility of a series R_t can be approximated by its absolute value transformation $|R_t|$,[12] we can define the "co-volatility" between the series $R_{1,t}$ and $R_{2,t}$ by the quantity $\sqrt{|R_{1,t}R_{2,t}|}$. We use Robinson's (1995) semi-parametric discrete version of Whittle (1962) approximate ML estimator in the spectral domain. This estimator, suggested by Künsch (1987), is based on the mild assumption that the spectrum $f(\lambda)$ of a long-memory time series can be approximated in the neighborhood of the zero frequency as

$$f(\lambda) \sim G\lambda^{-2d} \quad \text{as } \lambda \to 0^+ \quad \text{where } G \text{ is a strictly positive finite constant} \tag{26}$$

The consequences of a misspecification of the functional form of the spectrum in the Whittle estimator are avoided with this local approximation. After concentrating in G, the estimator is given by:

$$\hat{d} = \arg\min_d \left\{ \ln\left(\frac{1}{m} \sum_{j=1}^{m} \frac{I(\lambda_j)}{\lambda_j^{-2d}} \right) - \frac{2d}{m} \sum_{j=1}^{m} \ln(\lambda_j) \right\} \tag{27}$$

where $I(\lambda_j)$ is the periodogram estimated for the range of Fourier frequencies $\lambda_j = \pi j/T, j = 1, \ldots, m \ll [T/2]$, the bandwidth parameter m tends to infinity with T, but more slowly since

$$\frac{1}{m} + \frac{m}{T} \to 0 \quad \text{as} \quad T \to \infty \tag{28}$$

Under appropriate conditions, which include the existence of a moving average representation, this estimator is root-m consistent, asymptotically normally distributed, with distribution independent of the value of d:

$$\sqrt{m}(\hat{d} - d) \sim N\left(0, \frac{1}{4}\right) \tag{29}$$

Henry and Robinson (1996) have derived the formula for a data driven optimal bandwidth m_{opt}. They assume that the spectrum takes the form $f(\lambda) = \left|1 - e^{i\lambda}\right|^{-2d} h(\lambda)$ as $\lambda \to 0^+$, where $h(\cdot)$ is twice continuously differentiable with $h(0) > 0$. They derive the optimal bandwidth m_{opt} by minimizing the mean squared error:

$$m_{opt} = \left(\frac{3n}{4\pi}\right)^{\frac{4}{5}} |E_2(d)|^{-\frac{2}{5}} \tag{30}$$

where, following Delgado and Robinson (1996), $E_2(d)$ is approximated by :

$$E_2(d) = \frac{h''(0)}{2h(0)} + \frac{1}{12}d \tag{31}$$

[12] See Granger and Ding (1995).

where $h(0)$ and $h''(0)$, the first and third coefficients of the asymptotic expansion of the spectrum, are estimated by the periodogram least squares regression:

$$I(\lambda_j) = \left|1 - e^{i\lambda_j}\right|^{-2d} \sum_{l=0}^{2} \frac{\lambda_j^l}{l!}\beta_l + \textbf{residuals} \tag{32}$$

for the band of frequencies λ_j, $j = 1, \ldots, \hat{m}^{(0)}$, where $\hat{m}^{(0)} = n^{\frac{4}{5}}$ is a first approximation of the bandwidth parameter. Henry and Robinson (1996) proposed the iterative procedure

$$\hat{d}^{(k)} = \arg\min_d \left\{ \ln\left(\frac{1}{\hat{m}^{(k)}} \sum_{j=1}^{\hat{m}^{(k)}} \frac{I(\lambda_j)}{\lambda_j^{-2d}} \right) - \frac{2d}{\hat{m}^{(k)}} \sum_{j=1}^{\hat{m}^{(k)}} \ln(\lambda_j) \right\} \tag{33}$$

$$\hat{m}^{(k+1)} = \left(\frac{3n}{4\pi}\right)^{\frac{4}{5}} \left|E_2\left(\hat{d}^{(k)}\right)\right|^{-\frac{2}{5}} \tag{34}$$

We estimate $\hat{d}$ for several values of the bandwidth parameter m, and we obtain:

Table 1: Estimation of the fractional degree of integration for the series of absolute returns on Pound-Dollar $|R_{1,t}|$, Deutschmark-Dollar $|R_{2,t}|$, $\sqrt{|R_{1,t}R_{2,t}|}$ for the period April 1979 - January 1997. m_{opt} denotes the optimal bandwidth. Asymptotic standard errors $(2\sqrt{m})^{-1}$ are between parentheses

m	$\|R_{1,t}\|$	$\|R_{2,t}\|$	$\sqrt{\|R_{1,t}R_{2,t}\|}$
$T/4$	0.2385 (0.0147)	0.2312 (0.0147)	0.2413 (0.0147)
$T/8$	0.3071 (0.0207)	0.3219 (0.0207)	0.3230 (0.0207)
$T/16$	0.4113 (0.0293)	0.4073 (0.0293)	0.4393 (0.0293)
m_{opt}	0.3459 (0.0323)	0.3431 (0.0325)	0.3575 (0.0324)

Alternatively, we can estimate the intensity of long-memory in the volatilities and co-volatility by using the semiparametric estimators proposed by Giraitis, Kokoszka, Leipus and Teyssière (2000), henceforth GKLT, for the long-memory volatility model proposed by Giraitis, Robinson and Surgailis (1999). These estimators, based on the KPSS and the centered KPSS statistic, are in spirit similar to the Hurst (1951) R/S "pox-plot" estimator, developed by Mandelbrot and his co-authors: the range of the partial sum process $S_t = \sum_{j=1}^{t}(Y_j - \bar{Y}_T)$ is replaced by its second moment and its variance.

The standard R/S method is based on the rescaled range statistic defined as $\hat{R}_T/\hat{S}_T$:

$$\frac{\hat{R}_T}{\hat{S}_T} = \left(\max_{1\le k\le T} S_k - \min_{1\le k\le T} S_k\right)/\hat{S}_T \quad \text{where} \quad \hat{S}_T^2 = T^{-1}\sum_{j=1}^{T}(Y_j - \bar{Y}_T)^2 \tag{35}$$

where $\hat{R}_T$ is the range and $\hat{S}_T^2$ is a standard variance estimator. Given that

$$\sum_{j=1}^{k}(Y_j - \bar{Y}_T) = \sum_{j=1}^{k}(Y_j - EY_j) - \frac{k}{T}\sum_{j=1}^{T}(Y_j - EY_j) \tag{36}$$

and

$$\frac{1}{T^{1/2+d}}\sum_{j=1}^{[Tt]}(Y_j - EY_j) \stackrel{D[0,1]}{\longrightarrow} C\, W_{1/2+d}(t) \tag{37}$$

where C is a positive constant, and $\stackrel{D[0,1]}{\longrightarrow}$ means weak convergence in the space $D[0,1]$ endowed with Skorokhod topology. Then

$$\frac{\hat{R}_T}{T^{1/2+d}} \stackrel{d}{\to} C\left[\max_{0\le t\le 1} W^0_{1/2+d}(t) - \min_{0\le t\le 1} W^0_{1/2+d}(t)\right], \tag{38}$$

$W^0_{1/2+d}(t)$ being the fractional Brownian bridge: $W^0_{1/2+d}(t) = W_{1/2+d}(t) - tW_{1/2+d}(1)$. It is equally easy to verify that $\hat{S}^2_T \to \mathrm{Var}Y_1$. Thus,

$$\frac{1}{T^{1/2+d}}\frac{\hat{R}_T}{\hat{S}_T} \stackrel{d}{\to} \frac{C\left[\max_{0\le t\le 1} W^0_{1/2+d}(t) - \min_{0\le t\le 1} W^0_{1/2+d}(t)\right]}{\mathrm{Var}(Y_1)^{1/2}} \tag{39}$$

Equation (39) constitutes a theoretical foundation for the R/S estimator. Taking logarithms of both sides yields the heuristic identity:

$$\log\left(\hat{R}_T/\hat{S}_T\right) \approx \left(\frac{1}{2}+d\right)\log T + \textbf{constant}, \quad \text{as } T\to\infty, \tag{40}$$

which can also be written as $\hat{d}_{R/S} - d = O_P(1/\log T)$ with $\hat{d}_{R/S} = \frac{\log(\hat{R}_T/\hat{S}_T)}{\log T} - \frac{1}{2}$. Thus, $1/2+d$ can be interpreted as the slope of a regression line of $\log(\hat{R}_T/\hat{S}_T)$ on $\log T$.

GKLT (2000) observed that any other continuous functional of the partial sum process can be used in place of the range in equation (35) and considered the KPSS and centered KPSS statistics, respectively proposed by Kwiatkowski, Phillips, Schmidt and Shin (1992) and Giraitis, Kokoszka and Leipus (1999).

The KPSS statistic is defined as:

$$\hat{Z}_T = \frac{\hat{M}_T}{T\hat{S}^2_T} \quad \text{with} \quad \hat{M}_T = \frac{1}{T}\sum_{k=1}^{T} S_k^2. \tag{41}$$

i.e., the range has been replaced by the second moment. By equation (37),

$$\frac{\hat{M}_T}{T^{1+2d}} \stackrel{d}{\to} C^2\int_0^1 \left[W^0_{1/2+d}(t)\right]^2 dt. \tag{42}$$

Hence, defining $\hat{d}_{KPSS} = \log\hat{Z}_T/2\log T$, we get

$$\hat{d}_{KPSS} - d = O_P(1/\log T). \tag{43}$$

Thus, the slope of the regression line of $\log\hat{Z}_T$ on $\log T$ estimates $2d$.

Giraitis, Kokoszka and Leipus (1999) proposed a centering of the KPSS and defined the statistic $\hat{U}_T$ statistic

$$\hat{U}_T = \frac{\hat{V}_T}{\hat{S}^2_T T}, \tag{44}$$

where

$$\hat{V}_T = \frac{1}{T}\left[\sum_{k=1}^{T} S_k^2 - \frac{1}{T}\left(\sum_{k=1}^{T} S_k\right)^2\right]. \tag{45}$$

The range has been here replaced by the variance. As above, we conclude that the regression of $\log \hat{U}_T$ on $\log T$ estimates $2d$. Setting $\hat{d}_{V/S} = \log \hat{U}_T / 2\log T$, we get $\hat{d}_{V/S} - d = O_P(1/\log T))$.

The technical details of the implementation of these "pox-plot" estimators are described in Beran (1994) and GKLT (2000). The sample of T observations is subdivided in B adjacent and non-overlapping blocks of observations of equal size $[T/B]$. We then obtain a grid $t_1 = 1, t_2 = [T/B]+1, \ldots, t_i = (i-1)[T/B]+1, \ldots, t_B = T - [T/B] + 1$. For each point of the sequence $\{t_i\}_{i=1}^{B}$ we define a sequence of K increasing nested blocks with origin t_i, i.e., $\{[t_i, t_i + k_j]\}_{j=1}^{K}$, such that $t_i + k_j \leq T$, the sequence of K steps $\{k_j\}_{j=1}^{K}$ is given by a logarithmic grid. We calculate the R/S, V/S and KPSS statistics for each interval $\{\{[t_i, t_i + k_j]\}_{i=1}^{B}\}_{j=1}^{K}$ and obtain the sequences $\{\{R/S(t_i, k_j)\}_{i=1}^{B}\}_{j=1}^{K}$, $\{\{V/S(t_i, k_j)\}_{i=1}^{B}\}_{j=1}^{K}$, and $\{\{KPSS(t_i, k_j)\}_{i=1}^{B}\}_{j=1}^{K}$. We plot the logarithm of these statistics $\log(R/S(t_i, k_j)$, $\log(V/S(t_i, k_j)$, $\log(KPSS(t_i, k_j)$, against $\log(k_j)$ and then obtain a "pox-plot". We apply these "pox-plot" estimators on the series of volatilities and co-volatilities, and we obtain:

Table 2: Estimation of the fractional degree of integration for the series of absolute returns on Pound-Dollar $|R_{1,t}|$, Deutschmark-Dollar $|R_{2,t}|$, $\sqrt{|R_{1,t}R_{2,t}|}$ for the period April 1979 - January 1997.

	$\lvert R_{1,t}\rvert$	$\lvert R_{2,t}\rvert$	$\sqrt{\lvert R_{1,t}R_{2,t}\rvert}$
KPSS	0.2606	0.2340	0.2762
V/S	0.2473	0.2216	0.2533
R/S	0.2671	0.2350	0.2712

These results show that the two volatilities and the co-volatility share the same degree of long-memory for all the values of the bandwidth parameter m of the Whittle estimator, and for all the "pox-plot" estimators. Since the CCC model cannot account for the long-memory property of the conditional covariance, we expect that the unrestricted model, which takes into account the long-memory component in the conditional covariance, will have a better fit with the data.

3.2 Estimation of a bivariate model

We estimated the following bivariate long-memory ARCH:[13]

[13]We truncate the expansion of the infinite order polynomial at the order 2000.

$$\begin{pmatrix} R_{1,t} \\ R_{2,t} \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix} \quad \text{with} \quad \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} s_{11,t} & s_{12,t} \\ s_{12,t} & s_{22,t} \end{pmatrix}\right] \tag{46}$$

where $R_{1,t}$ and $R_{2,t}$ respectively denote the log of returns on the Pound-Dollar and Deutschmark-Dollar. We consider the five alternative parameterizations for the conditional covariance matrix

$$\textbf{(A)} \begin{pmatrix} s_{11,t} \\ s_{22,t} \\ s_{12,t} \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{\infty} \frac{B(p_1+j-1,d_1+1)}{B(p_1,d_1)} \varepsilon_{1,t-j}^2 \\ \sum_{j=1}^{\infty} \frac{B(p_2+j-1,d_2+1)}{B(p_2,d_2)} \varepsilon_{2,t-j}^2 \\ \rho\sqrt{s_{11,t}}\sqrt{s_{22,t}} \end{pmatrix} \tag{47}$$

$$\textbf{(B)} \begin{pmatrix} \sigma_{11,t}^{\delta} \\ \sigma_{22,t}^{\delta} \\ s_{12,t} \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{\infty} \frac{B(p_1+j-1,d_1+1)}{B(p_1,d_1)} |\varepsilon_{1,t-j}|^{\delta} \\ \sum_{j=1}^{\infty} \frac{B(p_2+j-1,d_2+1)}{B(p_2,d_2)} |\varepsilon_{2,t-j}|^{\delta} \\ \rho\sigma_{11,t}\sigma_{22,t} \end{pmatrix} \tag{48}$$

$$\textbf{(C)} \begin{pmatrix} s_{11,t} \\ s_{22,t} \\ s_{12,t} \end{pmatrix} = \begin{pmatrix} \frac{\omega_1}{1-\beta_1(1)} + \left(1 - \frac{(1-\phi_1 L)(1-L)^{d_1}}{1-\beta_1 L}\right) \varepsilon_{1,t}^2 \\ \frac{\omega_2}{1-\beta_2(1)} + \left(1 - \frac{(1-\phi_2 L)(1-L)^{d_2}}{1-\beta_2 L}\right) \varepsilon_{2,t}^2 \\ \rho\sqrt{s_{11,t}}\sqrt{s_{22,t}} \end{pmatrix} \tag{49}$$

$$\textbf{(D)} \begin{pmatrix} s_{11,t} \\ s_{22,t} \\ s_{12,t} \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{\infty} \frac{B(p_1+j-1,d_1+1)}{B(p_1,d_1)} \varepsilon_{1,t-j}^2 \\ \sum_{j=1}^{\infty} \frac{B(p_2+j-1,d_2+1)}{B(p_2,d_2)} \varepsilon_{2,t-j}^2 \\ \sum_{j=1}^{\infty} \frac{B(p_3+j-1,d_3+1)}{B(p_3,d_3)} \varepsilon_{1,t-j}\varepsilon_{2,t-j} \end{pmatrix} \tag{50}$$

$$\textbf{(E)} \begin{pmatrix} s_{11,t} \\ s_{22,t} \\ s_{12,t} \end{pmatrix} = \begin{pmatrix} \frac{\omega_1}{1-\beta_1(1)} + \left(1 - \frac{(1-\phi_1 L)(1-L)^{d_1}}{1-\beta_1 L}\right) \varepsilon_{1,t}^2 \\ \frac{\omega_2}{1-\beta_2(1)} + \left(1 - \frac{(1-\phi_2 L)(1-L)^{d_2}}{1-\beta_2 L}\right) \varepsilon_{2,t}^2 \\ \frac{\omega_3}{1-\beta_3(1)} + \left(1 - \frac{(1-\phi_3 L)(1-L)^{d_3}}{1-\beta_3 L}\right) \varepsilon_{1,t}\varepsilon_{2,t} \end{pmatrix} \tag{51}$$

Table 3 p 19 reports the QML estimates of the constant conditional variance long-memory ARCH model **(A)** for (*i*) the period 1979-1997, (*ii*) the same period subject to the constraint that the parameters of the two conditional variances are equal, i.e., $d_1 = d_2$ and $p_1 = p_2$, and (*iii*) for the period 1971-1997.

It clearly appears that for the period 1979-1997, the two series share the same degree of long-memory in the conditional variance, since the estimated values for d_1 and d_2 are very close. The constraints $d_1 = d_2$ and $p_1 = p_2$ are accepted by the likelihood ratio (LR) test, and the more parsimonious constrained model is then preferable. The estimation results for the period 1971-1997 are less appealing. The two long-memory parameters largely differ. This result completes Bollerslev's (1990) result, who observed with a multivariate GARCH model on weekly data, a difference between these two periods in the estimated parameters explained by the inception of the EMS in March 1979. The EMS system has modified both the short-term and the long-term variations of the volatilities of exchange rates returns.[14]

Table 4 p 19 reports the QML estimates of the constant conditional variance long-memory ARCH model, model **(B)**, for (*i*) the period 1979-1997, (*ii*) the same period with the restrictions $p_1 = p_2$ and $d_1 = d_2$, and (*iii*) for the period 1971-1997. The restriction $\delta = 2$ is rejected by the LR test. However, some parameters are instable for the two larger samples: this may be due to the misspecification of the model, as the assumption of a constant conditional correlation is too restrictive.

Table 5 p 20 displays the estimation results for the bivariate CCC–FIGARCH model. The log-likelihood value is higher than the one of the CCC long-memory ARCH model considered above, and the trade-off between the increase of the value of the log-likelihood function and the increase of the number of parameters leads us to choose the less parsimonious CCC–FIGARCH model.[15] As observed above, the degrees of long memory are the same for the period 1979-1997. Furthermore, the LR test accepts the hypothesis of equality of parameters β and ϕ for the two conditional variances. We will see later that the equality $\phi_1 = \phi_2$ is likely to be the consequence of the restrictive parameterization of the CCC–FIGARCH.

Tables 6 and 7 p 21 display the estimation results for respectively the unrestricted models **(D)**, and **(E)**. Column (*i*) contains the QML estimates for the unconstrained model, column (*ii*) contains the estimates for the model with the restrictions $d_1 = d_2 = d_3$, and column (*iii*) contains estimation results for the whole period 1971-1997. It is clear that the constant conditional correlation assumption for the models (**A, B, C**) is too strong since the value of the log-likelihood function is strongly improved. We also observe that the multivariate unrestricted FIGARCH model has a better fit with the data than the multivariate unrestricted long-memory ARCH model: the value of the log-likelihood

[14]We have also estimated the models **(A)** for the period 1971-1997, with two different conditional correlations, the first for the period 1971-1979, the second for the period 1979-1997. The LR test rejects the restriction that these two parameters are equal, and still rejects the restriction that the two degrees of long-memory are the same.

[15]The statistical properties of the standard parsimony criteria, the Akaike (AIC) and Bayes (BIC) Information Criteria, have not been established for the class of long-memory ARCH process. Beran, Bhansali and Ocker (1998) results on the BIC for ARFIMA$(p, d, 0)$ models casts doubts on the use of the BIC for long-memory volatility models.

function increases strongly, and usual parsimony criteria, e.g., the Akaike Information Criteria or the Bayes Information Criteria, favors the multivariate unrestricted FIGARCH model. As expected, the estimation of the multivariate FIGARCH is more problematic; a sensible choice of starting values is necessary, but the model looks well specified since we did not encounter too much problems. In this optimization problem, the conditional covariance matrix $\mathbf{\Sigma}_t$ should be positive definite for all t. Consequently, the step size of the line search procedure of the BFGS algorithm is quite small at the beginning of the optimization process, but the algorithm converges to an optimum satisfying the positivity constraint.

As expected from the results of the semiparametric analysis, all the elements of the conditional variance matrix, i.e., the two conditional variances and the conditional covariance, display the same degree of long-memory: the correlations between the estimated parameters are very high, and the restriction $d_1 = d_2 = d_3$ is accepted by the LR test for both models (**D**) and (**E**) for the EMS period. Furthermore, the restriction $\beta_1 = \beta_2 = \beta_3$ for model (**E**) is also accepted by the LR test for the same period. In a standard GARCH framework, the polynomial $\phi(L)$ captures the local variations, while the polynomial $\beta(L)$ captures the long-term variations. In our case, this interpretation is less obvious in a long-memory framework as the polynomial $\phi(L)$ enters in an intricate way in the expression of the infinite order lag polynomial $\alpha(L)$.

We estimate the models **(D, E)** with the restriction $d_1 = d_2 = d_3$ for the period 1971-1997. This restriction is accepted by the LR test. This result is due to the larger proportion of EMS observations in our sample, since this restriction is rejected when we consider only the first 5500 and 3500 observations from 1971.

This equality in the long-memory parameter for the same FX market suggests that this parameter is linked to the features of the market. A first consequence of the inception of a target zone system is a reduction of the level of variation of FX rates volatility. Interested readers are referred to Engle and Gau (1997) for more details. However, little more can be said using standard economic models. A theory has been recently produced by Kirman and Teyssière (2000), who considered a model with heterogeneous and interactive agents on FX markets. These agents are either "fundamentalists", i.e., they consider that FX rates do not depart too much from a series of fundamentals variables, or "chartists", i.e., they forecast FX rates by linear interpolation from the past. This model differ from standard "noise traders" models, by assuming that the size of each group is not fixed, but evolves through an epidemiologic process which herds on the extremes, i.e., the whole population is either fundamentalist or chartist. Kirman and Teyssière (2000) observed that (*i*) the level of long-memory is positively linked to the herding probability, and (*ii*) the property of common long-memory in the volatilities and co-volatility is obtained by assuming that the process governing the proportion of chartists/fundamentalists is the same for all markets.

4 Conclusion

We have proposed two classes of multivariate long-memory ARCH models. One of them, the constant conditional correlation (CCC) model, requires few constraints in the estimation procedure, but a strong restriction on the functional form of the conditional covariance matrix. This constraint appeared to be too strong since the long-memory CCC model is outperformed by the unrestricted long-memory conditional covariance model. We have applied these models to the bivariate modeling of the returns on two exchange rates, and have observed that for the period 1979-1997, after the EMS beginning, these two series have the same long-memory component in the conditional variance. The class of unrestricted conditional covariance models allowed us to go farther by showing that all the elements of the conditional covariance matrix share the same degree of long-memory.

This class of multivariate long-memory models can also be trivially extended by combining ARCH(∞) functions of different type in the conditional covariance matrix.

Another issue for future research is the extension of the concept of co–persistence in the conditional variance, developed by Bollerslev and Engle (1993), to the fractional case, although the idea of "fractional cointegration" for conditional variance processes is not yet supported by a theoretical model. (See Pagan (1996)).

References

ANDEL, J. (1986). Long-Memory Time Series Models. *Kybernetica*, **22,** 105–123.

BAILLIE, R. T., BOLLERSLEV, T. & H. O. MIKKELSEN (1996). Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, **74,** 3–30.

BERA, A. K. & M. L. HIGGINS (1993). On ARCH Models: Properties, Estimation and Testing. *Journal of Economic Surveys*, **7,** 305–366.

BERAN, J. (1994). *Statistics for Long-Memory Processes.* Chapman & Hall.

BERAN, J., BHANSALI, R. J. & D. OCKER (1998). On Unified Model Selection for Stationary and Nonstationary Short- and Long-Memory Autoregressive Processes. *Biometrika*, **85**, 921–934.

BOLLERSLEV, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, **31,** 307–327.

BOLLERSLEV, T. (1990). Modeling the Coherence in Short-run Nominal Exchange Rates: A Multivariate Generalized ARCH Model. *Review of Economics and Statistics*, **72,** 498–505.

BOLLERSLEV, T. & R. F. ENGLE (1993). Common Persistence in Conditional Variances. *Econometrica*, **61,** 166–187.

BOLLERSLEV, T., ENGLE, R. F. & D. B. NELSON (1994). ARCH Models. In R. F. Engle and D. McFadden editors, *Handbook of Econometrics*, **4**. North-Holland, Amsterdam.

BOLLERSLEV, T. & H. O. MIKKELSEN (1996). Modelling and Pricing Long-Memory in Stock Market Volatility. *Journal of Econometrics*, **73,** 151–184.

BOUGEROL, P. & N. PICARD (1992). Stationarity of GARCH Processes and of Some Nonnegative Time Series. *Journal of Econometrics*, **52l,** 115–127.

BREIDT, F. J., CRATO, N. & P. de LIMA (1998). On the Detection and Estimation of Long-Memory in Stochastic Volatility. *Journal of Econometrics*, **83,** 324–352.

DELGADO, M. A. & P. M. ROBINSON (1996). Optimal Spectral Bandwidth for Long-memory. *Statistica Sinica*, **6** , 97–112.

DING, Z. & C. W. J. GRANGER (1996). Modelling Volatility Persistence of Speculative Returns. A New Approach. *Journal of Econometrics*, **73,** 185–215.

DING, Z., GRANGER, C. W. J. & R. F. ENGLE (1993). A Long Memory Property of Stock Market Returns and a New Model. *Journal of Empirical Finance*, **1,** 83–106.

ENGLE, R. F. & G. G. J. LEE (1999). A Permanent and Transitory Component Model of Stock Return Volatility. In R. F. Engle and H. White editors, *Cointegration, Causality and Forecasting. A Festschrift in Honor of Clive W. J. Granger*, 475–497. Oxford University Press.

ENGLE, R. F. & K. F. KRONER (1995). Multivariate Simultaneous Generalized ARCH. *Econometric Theory*, **11,** 122–150.

ENGLE, R. F. & Y-F. GAU (1997). Conditional Volatility of Exchange Rates under a Target Zone. *University of California San Diego Discussion Paper* **97-06**.

ENGLE, R. F. (1982). Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, **50,** 987–1007.

GALLANT, A. R., HSU, C-T. & G. TAUCHEN (1998). Calibrating Volatility Diffusions and Extracting Integrated Volatility. In *Proceedings of the Second High Frequency Data in Finance Conference, HFDF-II,* , Olsen & Associates, 1–3 April 1998, Zurich, Switzerland.

GIRAITIS, L., KOKOSZKA, P. S. & R. LEIPUS (1999). Rescaled Variance and Related Tests for Long-memory in Volatility and Levels. *Mimeo LSE and University of Liverpool, Department of Mathematics.*

GIRAITIS, L., ROBINSON, P. M. & D. SURGAILIS (1999). A Model for Long-Memory Conditional Heteroskedasticity. *Annals of Applied Probability*, forthcoming.

GIRAITIS, L., KOKOSZKA, P. S. & R. LEIPUS (2000). Stationary ARCH Models: Dependence Structure and Central Limit Theorem. *Econometric Theory*, **16** , 3–22.

GIRAITIS, L., KOKOSZKA, P. S., LEIPUS, R. & G. TEYSSIÈRE (2000). Semiparametric Estimation of the Intensity of Long-memory in Conditional Heteroskedasticity. *Statistical Inference for Stochastic Processes*, **3,** 113–128.
`http://durandal.cnrs-mrs.fr/GREQAM/wp-pdf99/99A24.pdf`

GRANGER, C. W. J. & R. JOYEUX (1980). An Introduction to Long-memory Time Series Models and Fractional Differencing. *Journal of Time Series Analysis*, **1,** 15–29.

GRANGER C. W. J. & Z. DING (1995). Some Properties of Absolute Returns, an Alternative Measure of Risk. *Annales d'Économie et de Statistique*, **40,** 67–91.

GRAY, H. L., ZHANG, N. F. & W. A. WOODWARD (1989). On Generalized Fractional Processes. *Journal of Time Series Analysis*, **10,** 233-257.

MÄRDLE, W., SPOKOINY, V. & G. TEYSSIÈRE (2000). Adaptive Estimation of a Stochastic Volatility Model. *Humboldt University, Berlin, SFB 373 Discussion Paper* **00006.**

HARVEY, A. C. (1998). Long-Memory in Stochastic Volatility. In J. Knight and S. Satchell editors, *Forecasting Volatility in Financial Markets*, 307–320. Oxford Butterworth-Heineman.

HENRY, M. & R. PAYNE (1997). An Investigation of the Long Range Dependence in Intra-Day Foreign Exchange Rate Volatility. *LSE Financial Market Group Discussion Paper* **264**.

HENRY, M. & P. M. ROBINSON (1996). Bandwidth Choice in Gaussian Semiparametric Estimation of Long Range Dependence. In P. M. Robinson and M. Rosenblatt editors, *Athens Conference on Applied Probability and Time Series Analysis, Time Series Analysis, In Memory of E. J. Hannan*, **2**, 220–232. LNS 115. Springer Verlag, New York.

HORVÁTH, L. (2000). Change-Point Detection in Long-Memory Processes. *Journal of Multivariate Analysis*, forthcoming.

HOSKING, J. R. M. (1981). Fractional Differencing. *Biometrika*, **68**, 165–176.

HURST, H. E. (1951). Long-Term Storage Capacity of Reservoirs. *Transactions of the American Society of Civil Engineers*, **116**, 770–799.

KAZAKEVIČIUS, V. & R. LEIPUS (2000). On Stationarity in the ARCH(∞) Model. *Econometric Theory*, forthcoming.

KAZAKEVIČIUS, V., LEIPUS, R. & M-C. VIANO (2000). Stability of Random Coefficients Autoregressive Conditionally Heteroskedastic Models. *Mimeo Vilnius University.*

KIRMAN, A. P. & G. TEYSSIÈRE (2000). Microeconomic Models for Long-Memory in the Volatility of Financial Time Series. *GREQAM DT* **00A31.**
`http://durandal.cnrs-mrs.fr/GREQAM/wp-pdf00/00A31.pdf`

KOKOSZKA, P. S. & R. LEIPUS (2000). Change-Point Estimation in ARCH Models. *Bernoulli*, **6**, 513–539.

KÜNSCH, H. R. (1987). Statistical Aspects of Self-Similar Processes In Yu. Prohorov and V.V. Sazanov editors, *Proceedings of the First World Congress of the Bernoulli Society*, **1**, 67–74. VNU Science Press, Utrecht.

KWIATKOWSKI, D., PHILLIPS, P. C. B., SCHMIDT, P. & Y. SHIN (1992). Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root: How Sure Are We that Economic Series Have a Unit Root? *Journal of Econometrics*, **54**, 159–178.

MANDELBROT, B. B. & J. VAN NESS (1968). Fractional Brownian Motion, Fractional Noises and Applications. *S.I.A.M Review*, **10**, 422–437.

MANDELBROT, B. B. & J. M. WALLIS (1969). Computer Experiments with Fractional Gaussian Noises. *Water Resources Research*, **5**, 228–267.

MANDELBROT, B. B. & M. S. TAQQU (1979). Robust R/S Analysis of Long-run Serial Correlation. In *42nd Session of the International Statistical Institute, Manila, Book* **2**, 69–99.

MANDELBROT, B. B., FISHER, A. & L. CALVET (1997). A Multifractal Model of Asset Returns. *Cowles Foundation Discussion Paper* **1164**.

NELSON, D. B. (1990). Stationarity and Persistence in the GARCH(1,1) Model. *Econometric Theory*, **6**, 318-334.

PAGAN, A. R. (1996). The Econometrics of Financial Markets. *Journal of Empirical Finance*, **3**, 15–102.

RAY, B. K. & R. S. TSAY (2000). Common Long-range Dependence in Daily Stock Volatilities. *Journal of Business and Economics Statistics*, **18**, 254–262.

ROBINSON, P. M. (1991). Testing for Strong Serial Correlation and Dynamic Conditional Heteroskedasticity. *Journal of Econometrics*, **47**, 67–84.

ROBINSON, P. M. (1994). Time Series with Strong Dependence. In C.A. Sims editor, *Advances in Econometrics, Sixth World Congress*, 47–95. Cambridge University Press.

ROBINSON, P. M. (1995). Gaussian Semiparametric Estimation of Long Range Dependence. *The Annals of Statistics*, **23,** 1630–1661.

ROBINSON, P. M. & P. ZAFFARONI (1997). Modelling Nonlinearity and Long-Memory in Time Series. *Fields Institute Communications*, **11,** 161–170.

SPOKOINY, V. (1998). Estimation of a Function with Discontinuities via Local Polynomial Fit with an Adaptive Window Choice. *The Annals of Statistics*, **26,** 38-76.

WHITTLE, P. (1962). Gaussian Estimation in Stationary Time Series. *Bulletin of the International Statistical Institute*, **39,** 337-343.

A Estimation results

Table 3: Estimation results of bivariate long-memory ARCH process, with constant conditional correlation, model (**A**). (Robust t-statistics between parentheses).

	(i) 1979-1997	(ii) 1979-1997	(iii) 1971-1997
μ_1	-0.0032 (-0.32)	-0.0036 (-0.30)	-0.0072 (0.80)
μ_2	-0.0003 (-0.03)	0.0000 (0.00)	-0.0174 (1.41)
p_1	5.8472 (4.42)	6.0607 (5.66)	2.9486 (2.67)
p_2	6.3221 (4.88)	—	4.1063 (2.50)
d_1	0.4496 (8.11)	0.4379 (11.66)	0.5495 (3.916)
d_2	0.4309 (11.47)	—	0.3475 (14.70)
ρ	0.7196 (67.49)	0.7197 (67.54)	0.6192 (32.61)
Log-likelihood	-7139.2294	-7139.7738	-9388.7127

Table 4: Estimation results of bivariate long-memory ARCH process, with constant conditional correlation, model (**B**). (Robust t-statistics between parentheses).

	(i) 1979-1997	(ii) 1979-1997	(iii) 1971-1997
μ_1	-0.0047 (0.52)	-0.0043 (-0.49)	-0.0066 (0.89)
μ_2	-0.0018 (0.17)	-0.0012 (-0.12)	-0.0096 (1.68)
p_1	2.5931 (3.18)	2.8592 (3.77)	2.3645 (1.69)
p_2	3.1258 (3.61)	—	1.8856 (3.24)
d_1	0.2385 (3.98)	0.2413 (4.60)	0.6657 (2.48)
d_2	0.2427 (4.73)	—	0.2192 (4.12)
δ	2.3687 (19.49)	2.3661 (20.12)	2.4004 (13.08)
ρ	0.7383 (72.86)	0.7380 (72.89)	0.6628 (41.43)
Log-likelihood	-7117.6135	-7119.7266	-9368.2204

Table 5: Estimation results of bivariate CCC–FIGARCH, model (**C**). (Robust t-statistics between parentheses).

	(i) 1979-1997	(ii) 1979-1997	(iii) 1971-1997
μ_1	-0.0039 (-0.44)	-0.0036 (-0.41)	-0.0059 (-0.88)
μ_2	-0.0020 (-0.20)	-0.0014 (-0.14)	-0.0154 (-1.81)
ω_1	0.0218 (2.79)	0.0230 (3.15)	0.0081 (2.70)
ω_2	0.0320 (2.71)	0.0293 (3.18)	0.0441 (1.37)
β_1	0.5638 (7.27)	0.5628 (9.82)	0.7586 (9.28)
β_2	0.5573 (9.18)	—	0.4016 (3.00)
d_1	0.2882 (6.40)	0.2727 (6.67)	0.6692 (4.14)
d_2	0.2529 (5.27)	—	0.2302 (2.65)
ϕ_1	0.3687 (5.20)	0.3775 (7.56)	0.3010 (3.06)
ϕ_2	0.3862 (7.33)	—	0.2895 (3.50)
ρ	0.7448 (76.00)	0.7446 (75.58)	0.6734 (46.76)
Log-likelihood	-7090.8634	-7091.9627	-9198.5869

Table 6: Estimation results of bivariate long-memory ARCH process, model **(D)**. (Robust t-statistics between parentheses).

	(i) 1979-1997	(ii) 1979-1997	(iii) 1971-1997
μ_1	0.0008 (0.09)	0.0008 (0.09)	-0.0078 (-0.91)
μ_2	0.0007 (0.07)	0.0007 (0.07)	-0.0067 (-0.84)
p_1	4.7013 (3.82)	4.6433 (4.67)	4.1442 (3.19)
p_2	5.1265 (4.72)	5.1192 (5.01)	3.7340 (2.89)
p_3	5.5434 (4.90)	5.4724 (5.91)	3.8172 (4.29)
d_1	0.4683 (5.89)	0.4655 (6.89)	0.3817 (6.18)
d_2	0.4661 (6.24)	—	0.3648 (8.35)
d_3	0.4705 (5.20)	—	0.3526 (5.42)
Log-likelihood	-6736.0071	-6736.6535	-8771.3637

Table 7: Estimation results of the unrestricted bivariate FIGARCH, model (**E**). (Robust t-statistics between parentheses).

	(i) 1979-1997	(ii) 1979-1997	(iii) 1971-1997
μ_1	-0.0011 (-0.13)	-0.0009 (-1.11)	-0.0052 (-0.88)
μ_2	-0.0008 (-0.90)	-0.0006 (-0.07)	-0.0063 (-0.83)
ω_1	0.0108 (2.84)	0.0107 (2.97)	0.0062 (2.00)
ω_2	0.0131 (2.91)	0.0129 (3.00)	0.0149 (1.10)
ω_3	0.0085 (2.43)	0.0085 (2.50)	0.0031 (0.61)
β_1	0.6417 (10.91)	0.6414 (12.82)	0.6895 (7.19)
β_2	0.6376 (11.68)	—	0.5987 (2.58)
β_3	0.6432 (11.15)	—	0.6388 (3.98)
d_1	0.4235 (7.12)	0.4187 (7.75)	0.5650 (3.91)
d_2	0.4179 (6.02)	—	0.4278 (1.32)
d_3	0.4225 (6.88)	—	0.4741 (2.07)
ϕ_1	0.3243 (8.29)	0.3296 (10.24)	0.2799 (4.35)
ϕ_2	0.3089 (8.26)	0.3197 (9.75)	0.2860 (3.99)
ϕ_3	0.2965 (8.93)	0.2994 (9.60)	0.2649 (3.69)
Log-likelihood	-6683.9150	-6684.1682	-8586.5104