JOURNAL DE LA SOCIÉTÉ STATISTIQUE DE PARIS

IGNACIO MAULEÓN

Instability and long memory in conditional variances

Journal de la société statistique de Paris, tome 138, nº 4 (1997), p. 67-88
<http://www.numdam.org/item?id=JSFS_1997__138_4_67_0>

INSTABILITY AND LONG MEMORY IN CONDITIONAL VARIANCES

Ignacio MAULEÓN [1]
Universidad de Salamanca [2]

Abstract

The primary concern of this paper is to propose and test a new specification for, first, the conditional variance, intending to capture a long memory property, and second, the density function based on an Edgeworth expansion, of high frequency financial data. The specification is tested with twenty five years of daily observations from the main world financial markets, and the empirical results support it in most cases.

Keywords: unstable conditional variance, Edgeworth-Sargan density, long memory, high frequency financial observations.

JEL classification : C53, F31.

---

1. This paper was presented at the *Financial Market Dynamics* conference held at the Université d'Evry-Val d'Essonne (Paris, 1996). The author is specially grateful to the comments of one anonymous referee, while retaining responsability for any remaining error. The research of this paper has been financed by the DIGCYT under grant PB94-1502.

2. Departemento de Economía e Historia, Fac. de Economía y Empresa (edif. FES) Campus Miguel de Unamuno, E-37008 SALAMANCA (Spain). Phone : 34.23.294640 Fax : 34.23.294686 – Email : imaul@gugu.usal.es

1. Introduction

In empirical applications it is often found that conditional variances of the residuals of econometric models, are near or *non covariance stationary*, a typical example being the GARCH(1,1) with the sum of both coefficients being close to, or even above, unity (this is usually referred in the literature as *non covariance stationarity*, and it implies that the unconditional variance does not exist). This empirical property has been found, among others, by ENGLE and BOLLERSLEV [1986], DELONG and SUMMERS [1986], PAGAN and SCHWERT, [1990a, 1990b], and PHILLIPS and LORETAN, [1994, 1995]. The implications of this empirical regularity are, at least, twofold: **1)** from an economic point of view, it is hard to find a satisfactory theoretical explanation and, empirically, it is unlikely that the variance of the rate of change of an economic variable tends to infinity (see GOODHART and O'HARA [1996] for a more thorough discussion of this point); **2)** econometrically, it poses strong problems of estimation and statistical interpretation of results (for example, if the regressors are stationary, their impact on the dependent variable relatively to the error would decrease along time, which might lead to inconsistent estimation in some cases). One strand of econometric research has tried to address the problem directly, although the results are still applicable only to specific models (HARVEY and ROBINSON [1991], WOOLDRIDGE and WHITE [1988], DAVIDSON [1991], and HANSEN [1995]). While these results are extremely useful, and it may well be that for many stochastic volatility processes the *non covariance stationarity* is not spurious, in other situations their applicability might be less clear as pointed out previously. Another strand of research has focussed on misspecification as a likely cause for strong, and even *non covariance stationary* ARCH and GARCH processes (LASTRAPES [1989], LAMOUREUX and LASTRAPES [1990], and CAI [1994]). The research presented in this paper has an empirical content, primarily, and it follows this last line of work. A number of plausible misspecifications giving rise to strong ARCH effects are discussed first (section 2). These include changes in the constant of the conditional variance process and, specially, long, but otherwise stable, ARCH effects (namely, a long memory property of the conditional variance). A new approach to fit this type of model is suggested, based on a smooth polynomial that depends on a few parameters, to a *very long*, although truncated, ARCH process. An Edgeworth-Sargan type of density is suggested for the conditional distribution of the error. The empirical results obtained after fitting this combined specification to the major world financial series are presented in section 3 (daily observations spanning 25 years, for exchange rates, stock markets, and short term interest rates, for a selection of countries). A last section gathers the main empirical results, and some technical details are left to an appendix.

2. Analytical background

The problem of misspecification causing Arch effects, has been tackled by a number of authors. LAMOUREUX and LASTRAPES [1990], and LASTRAPES [1990], by means of simulations, show that a change in the mean of the ARCH process (i.e. a structural break), goes some way towards explaining a large ARCH root (the suggestion being similar to that of PERRON [1989], to explain a unit root in the mean of the dynamic process). This idea is pursued by CAI [1994], who applies the switching regime model proposed by HAMILTON [1994], and shows how an unstable GARCH(1,1) model simply vanishes, collapsing into a two state model with constant conditional variances within each regime (but different between regimes). However, CAI also introduces a change in the mean of the dependent variable, which is a likely explanation for the strong result obtained (since misspecifications of the mean of the dependent variable induce Arch effects, as it has long been recognized). A heuristic explanation of this result could be provided by considering, first, the random variable $x_t = u^2{}_t$, where u_t is given by,

$$\begin{aligned} u_t &= N(0,\sigma^2) \quad , \; t \leq T/2 \\ &= N(0,k.\sigma^2) \quad , \; t > T/2 \end{aligned} \tag{2.1}$$

It is now a matter of a few straightforward calculations to derive the next result,

$$\begin{aligned} &\text{plim}\,[\Sigma^T{}_1(x_t - x^*)\,.\,(x_{t-1} - x^*)]\,/\,\Sigma^T{}_1(x_t - x^*)^2] = \\ &\quad = (1-k)^2\,/\,\{(1-k)^2 + 4(1+k^2)\} \end{aligned} \tag{2.2}$$

(x^* being the sample mean of x_t: since this last expression is different from zero whenever k is not one, this suggests that a change in the variance unaccounted for, can yield ARCH effects, (LAMOUREUX and LASTRAPES [1990] show this to be the case by means of Monte Carlo simulations).

The second type of misspecification is related to the dynamics of the ARCH model itself. To start the discussion, consider the following ARCH model,

$$h_t = q_0 + \Sigma^p{}_1 q_s\,.\,u^2{}_{t-s} \tag{2.3}$$

where h_t is the conditional variance of u_t, and, $q_0 > 0$, $0 < \Sigma^p{}_1 q_s < 1$. It is not too hard to see that if a GARCH(1,1) is fitted to (2,3), a large autoregressive root might be obtained, this root trying to capture the long lag in the serial correlation. In fact, if the model $h_t = a\,.\,h_{t-1} + b\,.\,u^2{}_{t-1}$ is fitted by ordinary least squares (assuming the u's were known), the following result is obtained,

$$\text{plim}\,(a^* + b^*) = (p-1)\,/\,p + q \tag{2.4}$$

(where a^*, b^*, are the OLS estimates of a, b, and $q_s = q$, for the sake of clarity). From this last expression we see that if, for example, $p = 6$, and $q = 0.1$, the sum of the autoregressive coefficients in the ARCH model of (2.3) is 0.6, whereas the value given by (2.4) would be, 0.933 ($p = 12, q = 0.05$, would yield 0.6 and 0.97 as the corresponding values): that is, a fitted GARCH(1,1) model would be close to, or non covariance stationary, when the true model was far from that case (as it will be seen in the next section, this happens to be the case of some data analysed in this paper). Another aspect of the GARCH model is that it yields an inadequate lag shape for the conditional variance model (it decays too slowly at the beginning, and too quickly at long lags; this was noted by DING *et. al.* [1993]).

To summarize, misspecification of at least the following three kinds may cause large ARCH effects: **1)** changes in the conditional variance process (for example, in the constant); **2)** long, but otherwise stable, lags in the ARCH process, (i.e. a stable ARCH(p) with a large value for p, can be proxied by an unstable GARCH(1,1) model); **3)** inappropriate account of thick tails of the error distribution. To take account of these points, two suggestions are made in this paper: first, an Edgeworth-Sargan type of density; second, a new model to fit the long memory property of the ARCH process. The Edgeworth-Sargan density is given by the following specification (polynomials up to order eight are presented),

$$\begin{aligned}
g(\varepsilon_t) &= f(\varepsilon_t)\,.\,\{1+\Sigma^q{}_{s=3}\; d_s\,.\,H_s(\varepsilon_t)\}\\
H_3(\varepsilon_t) &= \varepsilon_t{}^3 - 3\,.\,\varepsilon_t\\
H_4(\varepsilon_t) &= \varepsilon_t{}^4 - 6\,.\,\varepsilon_t{}^2 + 3\\
H_5(\varepsilon_t) &= \varepsilon_t{}^5 - 10\,.\,\varepsilon_t{}^3 + 15\,.\,\varepsilon_t\\
H_6(\varepsilon_t) &= \varepsilon_t{}^6 - 15\,.\,\varepsilon_t{}^4 + 45\,.\,\varepsilon_t{}^2 - 15\\
H_7(\varepsilon_t) &= \varepsilon_t{}^7 - 21\,.\,\varepsilon_t{}^5 + 105\,.\,\varepsilon_t{}^3 - 105\,.\,\varepsilon_t\\
H_8(\varepsilon_t) &= \varepsilon_t{}^8 - 28\,.\,\varepsilon_t{}^6 + 210\,.\,\varepsilon_t{}^4 - 420\,.\,\varepsilon_t{}^2 + 105 \qquad (2.5)
\end{aligned}$$

It is easily checked that $E(\varepsilon_t) = 0$, $E(\varepsilon_t{}^2) = 1$. The value of d_3 accounts for asymmetry ($d_3 = E(\varepsilon_t{}^3)\,/\,6$), and d_4 for kurtosis ($d_4 = (E\varepsilon_t{}^4 - 3)\,/\,24$). Finally, $f(\varepsilon_t)$ stands for a $N(0,1)$ density. This density function has several advantages over its competitors, among the most obvious being the following two: **1)** it can be easily generalized to include more parameters, should they be needed, so that it can fit any type of density (GALLANT and NYCHKA [1987]), and, **2)** the probability distribution function is easily obtained (since $d^n f(.) = (-1)^n H_n(.) f(.)$, higher order polynomials $H_n(.)$ can be calculated easily; see, for example, KENDALL and STUART [1977]). Early theoretical applications in econometrics of this type of distribution can be found in SARGAN [1976], and MAULEÓN [1983]. Some empirical applications are presented in BOURGOIN and PRIEUL [1997], and MAULEÓN and PEROTE [1997].

The model proposed to fit a long, though truncated, ARCH process is presented next. It is based on the idea that a transformation of the lag process, leads to a set of coefficients that must lie on a smooth function of the lag length. It is suggested that this new set of coefficients could be approximated, for example, by an Almon polynomial (although other approximations are not ruled out). The model is given by the following set of equations ($\sigma^2{}_t = E[u^2{}_t\,/\,t-1]$):

$$\begin{aligned}
\sigma^2{}_t &= \phi_0 + \Sigma^N{}_1(\phi_s\,.\,u^2{}_{t-s}) \qquad \phi_s \geq 0\\
&= a_0 + a_1\,.\,u^2{}_{t-1} + \Sigma^n{}_{j=2}(a_j\,.\,w_j) \qquad (2.6)
\end{aligned}$$

where,

$$w_j = \Sigma_{s=1}{}^{N-1}[s^{j-2} \,.\, \Delta\,(u^2{}_{t-s})] \quad , \quad j = 2, \dots , n. \tag{2.7}$$

and the parameters ϕ_s are certain function of the parameters a_s (see the Appendix), and in particular,

$$a_0 = \phi_0$$
$$a_1 = \Sigma^N{}_1\, \phi_s \tag{2.8}$$

Typically, n will be much smaller than N, so that the scale of the optimization problem will be reduced. Also, a_1 measures the accumulated impact of past squared errors, so that it must be less than one for covariance stationarity (since the unconditional variance is obtained taking unconditional expectations in (2.6), $\sigma^2 = a_0 + a_1\sigma^2$, so that $\sigma^2 = a_0\,/\,(1 - a_1)$, provided $a_1 < 1$).

Altogether, the two set of equations (2.5,6) provide a fairly flexible structure, to model jointly the ARCH process, and the distribution of the error (skewness, kurtosis, and higher order departures from Normality considered; an alternative is proposed by NELSON [1991], but perhaps it is more complicated to implement computationally). In order to combine both properties, that is, the non normality and the conditional heteroskedasticity, it is convenient to define the variable u_t, as $u_t = \sigma_t \,.\, \varepsilon_t$, so that its density function $\phi\,(.)$ will be given by,

$$\phi(u_1) = g(u_t\,/\,\sigma_t)\,/\,\sigma_t \tag{2.9}$$

where $g\,(.)$ is the ES density of (2.5). Empirical results based on this model are presented in the next section.

3. Empirical results

The data set used to implement the model of the previous section is fairly large: daily observations for the main financial assets are considered (interest rates, exchange rates and stock market indices), for the largest world economies (USA, Japan, Germany, France and the UK), with a sample period spanning from the beginning of 1971 to 1996 (eleven individual series, with almost seven thousand observations some of them, altogether). The reason behind the choice of such a wide data set was to provide some general validity to the model proposed in this paper, in case the empirical results supported it – as it will be seen that they do, in this section.

The empirical results are given next. The symbols in the tables match those of the sets of equations given in the previous section, and LK is the value of the likelihood in logs. (this could be compared with the fit yielded by other specifications in order to select among non nested models). In all cases, the variable under study is the corresponding daily rate of change. First, Table 1 presents a descriptive analysis of covariance stationarity by comparing the sum of all coefficients of the *long* ARCH (the first equation in (2.6), i.e., $a_1 = \Sigma^{\infty}1\,\phi_s$), to the covariance stationarity condition of a GARCH(1,1) process (where the parameters (α, B) are defined by the expression, $\sigma^2{}_t =$

$\phi_0 + \alpha \cdot u^2_{t-1} + B\sigma^2_{t-1}$). Overall, the results support the conclusion that the series are covariance stationary, although the GARCH model may give another impression (this is particularly true for the Japanese and USA stock markets). Further evidence is provided in Fig. 4 in the Appendix, where a moving average based on one thousand observations, of the squared rate of changes of the Nikkei index is presented : although the series does not display constancy along time, and it might even be possible to argue about dependency on time of the conditional variance, it would be difficult to argue that it tends to infinity.

Table 1 – **Covariance stationarity analysis**

Results for the Stock Market.

	France	Germany	Japan	UK	USA
Garch α+B	.97	.96	1.03	.985	1.0
a_1	.868	.81	.885	.883	.796

Results for Exchange Rates.

	France	Germany	Japan	USA
Garch α+B	.97	.96	.97	.98
a_1	.9016	.829	.921	.871

Results for Short Term Interest Rates.

	Japan	USA
Garch α+B	.92	.97
a_1	.848	.845

The complete results for the models fitted to the stock market are presented in Table 2. The proposed specification fits the data acceptably and, as noted before, all series are covariance stationary, since the estimated value for a_1 is less than unity and, therefore, the unconditional variance is finite, as noted in the previous section (according to the fitted model, at least). The length of the lag of the ARCH process, has been selected setting an upper limit of 200 lags (near one year, taking account of weekends) initially : shorter lags were tested and accepted in all cases (around 35, equivalent to 7 weeks, is the value accepted in most cases). Therefore, from this point of view it can also be concluded that the persistence of volatility is limited. The Almon polynomial has six parameters in all cases $(a_2, \ldots, a_7)$, which are generally significant. The value of the polynomial was set to six for practical reasons : it was considered that this would provide enough flexibility and, on the other hand, higher values might lead to collinearity of the $w's$, complicating the non-linear estimation excessively (see the Appendix for further discussion). A likelihood ratio test for the joint significance of all a_s2's is also provided (denoted by $LKR_{(a's)}$) : this test is asymptotically distributed as Chi-squared with 6 degrees of freedom, so that the null is always strongly rejected (the 95 % value is 12.6). No asymmetric effect was found in the density of the errors of the model (all odd polynomials were not significant). The specification performs acceptably, as well, for exchange rates (see Table 3 in the Appendix), although the sample had to be adjusted in some cases (the exchange rates are measured in domestic units per British pounds). In the case of short term interest rates, however, the model only works satisfactorily, to some extent, for the Japanese and USA cases (see Table 4 in the Appendix). For the European markets, the specification yielded systematically values of the sum of all lagged ARCH effects, well above unity, failing to improve, therefore, the GARCH specification (which also yielded values of $(\alpha + B)$ above one). This may be interpreted according to the results of Lamoureux and Lastrapes, and the derivation given in (2.1,2) in the previous section : the reason may be that short term interest rates are a policy variable of the monetary authorities and, therefore, manipulated to a large extent (this is likely to cause frequent and sudden changes in both, their value and volatility and, therefore, to induce spurious ARCH effects). Finally, it is reasonable to assume that this problem has affected European rates primarily, because of the need to stabilize exchange rates inside the bands of the EMS.

Table 2 – **Estimation results for the stock market**

	France	Germany	Japan	UK	USA
a_0	.345E-4 (18.8)	.19E-4 (7.8)	.176E-4 (13.2)	.113E-4 (6.1)	.178E-4 (7.3)
a_1	.834 (32.3)	.81 (19.2)	.885 (28.8)	.883 (23.4)	.796 (18.1)
a_2	-.784 (29.1)	-.826 (17.5)	-.952 (24.2)	-.86 (20.2)	-.795 (16.9)
a_3	.092 (13.0)	.12 (10.6)	.075 (4.5)	.081 (7.0)	.0091 (.5)
a_4	-.0039 (8.2)	-.0086 (6.7)	-.0033 (1.2)	-.00394 (3.1)	.0059 (1.9)
a_5	-.34E-5 (40.3)	.00027 (4.4)	-.144E-4 (.08)	.00011 (1.9)	-.00043 (1.9)
a_6	.366E-5 (5.6)	-.38E-5 (2.9)	.303E-5 (.57)	-.166E-5 (1.6)	.113E-4 (1.6)
a_7	-.580E-7 (4.8)	.184E-7 (1.9)	-.293E-7 (.5)	.106E-7 (1.4)	-.103E-6 (1.3)
d_3	n.s..	n.s.	n.s.	n.s.	n.s.
d_4	.126 (15.7)	.0455 (6.8)	.184 (25.3)	.031 (5.4)	.058 (8.7)
d_6	.0224 (10.0)	.0082 (4.7)	.0324 (16.0)	.0066 (4.2)	.0088 (4.7)
d_8	.00263 (11.9)	.00066 (3.6)	.0042 (21.5)	.0006 (3.5)	.00105 (5.4)
T	5971 24/02/71 12/01/94	6590 24/02/71 28/05/96	6590 24/02/71 28/05/96	6590 24/02/71 28/05/96	6590 24/02/71 28/05/96
LK	19638.8	22131.4	22847.7	22236.8	22240
$LKR_{(a's)}$	605	882	1164	1196	668
N	35	50	35	35	55

The behaviour of the Edgeworth-Sargan density is examined in Fig. 1, where a comparison to the empirical (after fitting the model) conditional density is carried out : the Normal density performs poorly, as expected, and the ES density behaves acceptably, although there is a peak at the middle, unaccounted for (the results correspond to the Japanese model, which is the case that yields the highest values for the non-Normal coefficients ; the d_s's of (2.5)). The behaviour of the various approximations at the tails, is further examined graphically in the Appendix (see Fig. 3, in the appendix). A formal *goodness of fit* test along the lines discussed in KENDALL and STUART [1977] is also discussed in the Appendix.

An example of the shape of the lag coefficients of squared residuals on the conditional variance, produced by the model suggested in the previous section, is presented in Fig. 2 (denoted LARCH), to stand for *Long* ARCH, and compared to the corresponding lags generated by a GARCH process (results for the short term Japanese interest rate) : the GARCH specification totally misses the initial peak at lag one, decreases too slowly at the beginning and, finally, decreases too quickly, so that it does not pick the long tailed effect either (this empirical regularity has been noted by other researchers, as well ; see again, for instance, DING *et al.* [1993]).

Fig. 1 – **Comparison of densities**

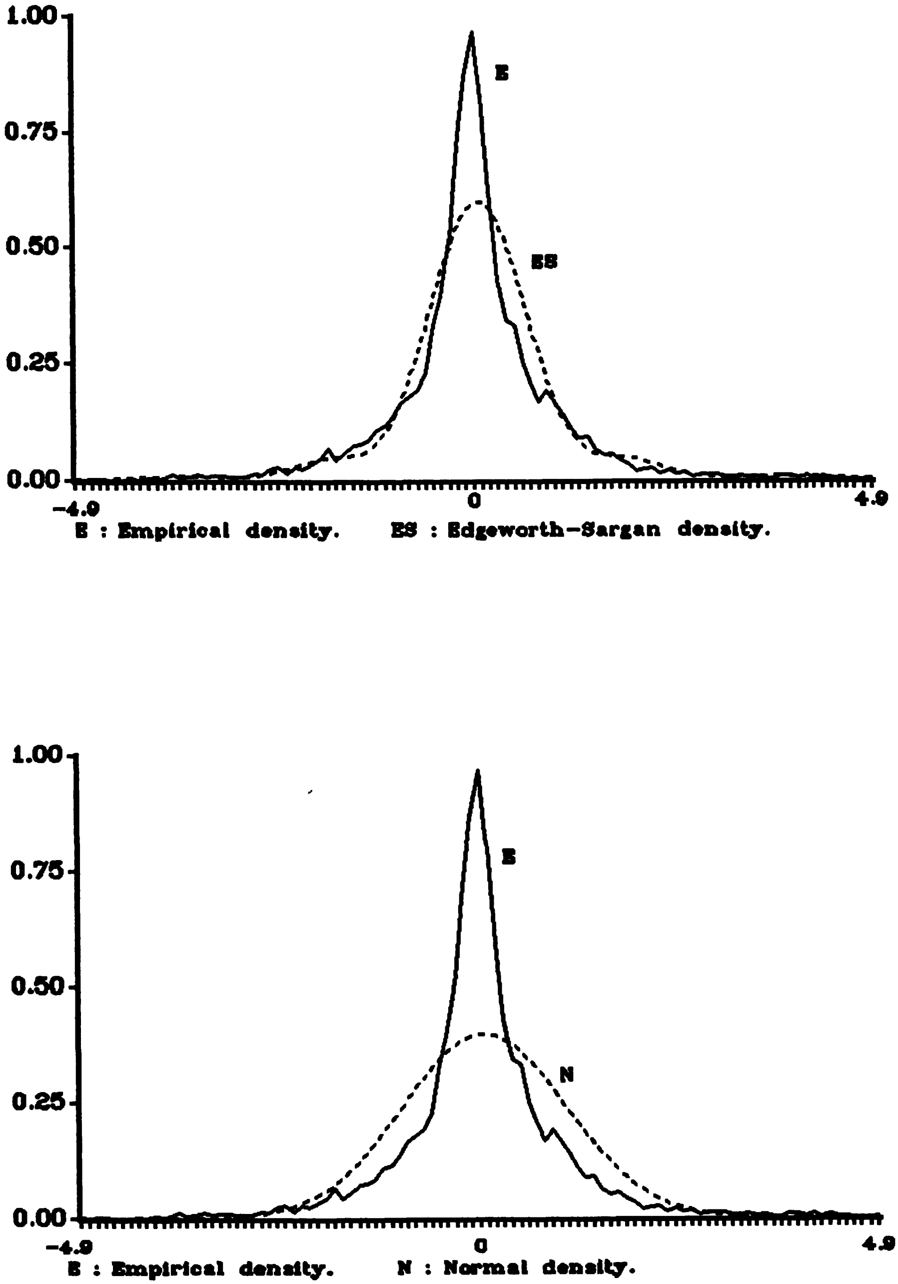
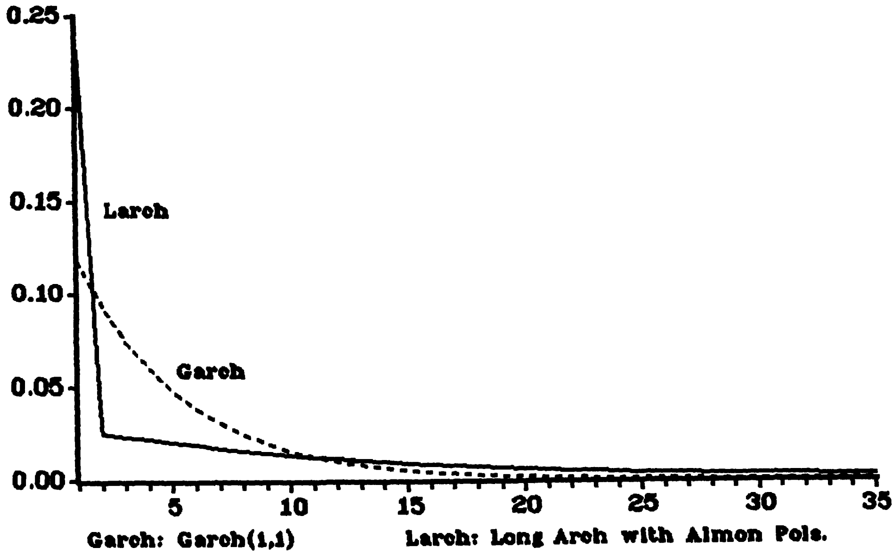
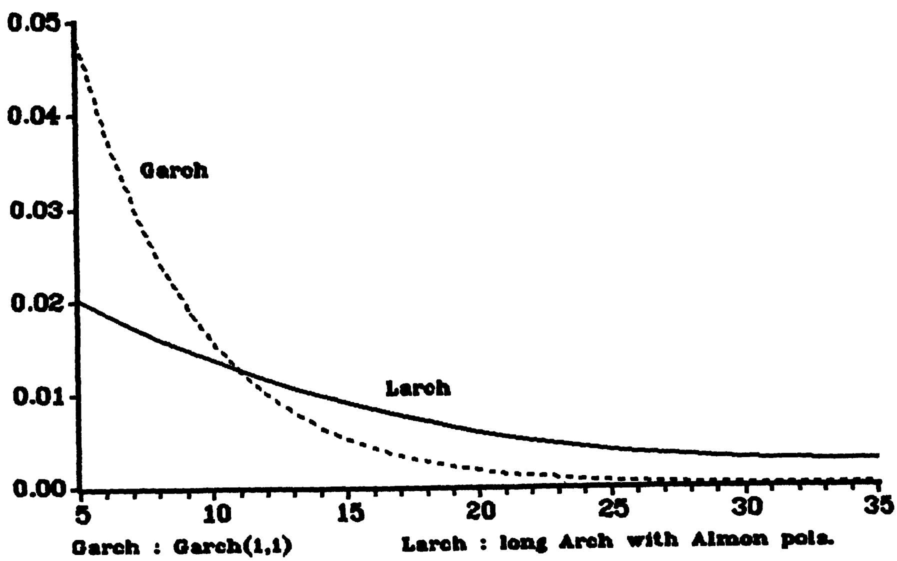


Fig. 2 – Conditional variance dynamics





4. Summary and conclusions

The main focus of the research reported in this paper is empirical. The purpose has been to suggest a new model to take into account the long memory property of conditional variances, but avoiding the complications involved in other procedures (for example, the fractional integrated process; see SOWELL [1992]). Besides, a new type of density has been suggested, based on Edgeworth expansions. Both specifications have been fitted to several financial series, and the empirical results, on the whole, support them (although this is not always the case). Other results are the following : **1)** it has been shown that a change in the conditional variance may yield some, though limited, ARCH effects. This might explain why, in the case of some monetary policy variables, the fitted ARCH, or GARCH processes, are unstable ; **2)** it has been shown that an unstable GARCH(1,1) process can be a spurious result, pointing to an underlying stable ARCH(p) process with a long lag, *p*. This has been found to be the case in several of the estimated models presented in the paper.

Although the testing of the specification suggested in this paper has been rather extensive, other cases could be covered in future research (long term interest rates, for example, and other exchange rates). Comparisons of the ES density with other densities, and alternatives to the Almon polynomials to fit the long ARCH lags, would be two of the most immediate extensions of the research presented in this paper.

Appendix

A.1. The optimization methodology and further empirical results

First, the results for exchange rates and short term interest rates are given (Tables 3 and 4). It should be noted that, in the case of the short term rates, it was only possible to estimate a density with two polynomials : adding further terms, produced very significant, but unacceptable results, as the implied densities had some negative values. Then, the detail of the behaviour at the tails of the different densities considered in the paper is offered in Fig. 3. Finally, it may be convenient to make a few comments on the optimization algorithm employed in the paper. Since the model is highly non linear, the optimization technique is not straightforward : in fact, unless the starting values are chosen carefully, the algorithm does not converge at all, or it does so at a local minimum. The strategy implemented in this research involved iterating between the density, and the long ARCH parameters. In every step, the optimization was carried out in two additional steps : first, an algorithm based on analytical first order derivatives, that used the covariance of the analytic gradient as an approximation to the Hessian in the iteration, was implemented; second, and once the optimization was supposed to be reasonably close to the global optimum, a full Newton-Raphson procedure with analytical first and second order derivatives was implemented (under the assumption that, near the optimum, the likelihood must be approximately quadratic).

In every step the parameters were added in a stepwise fashion, rather than being estimated jointly in a first step. For instance, in order to estimate the parameters of the density (the d_n's), the starting point was a $N(0, \sigma)$ density with constant variance; then, d_4 was added and estimated, and so forth for the remaining d_n coefficients. Initial values for the d_n's are provided by the following results,

$$
\begin{aligned}
d_3 &= \mu_6 \,/\, 6 \\
d_4 &= (\mu_4 - 3) \,/\, 24 \\
d_5 &= (\mu_5 - 10 \,.\, \mu_3) \,/\, 120 \\
d_6 &= (\mu_6 - 15 \,.\, \mu_4 + 30) \,/\, 720 \\
d_7 &= (\mu_7 - 21 \,.\, \mu_5 + 105 \,.\, \mu_3) \,/\, 5040 \\
d_8 &= (\mu_8 - 28 \,.\, \mu_6 + 210 \,.\, \mu_4 - 315) \,/\, 40320
\end{aligned}
\tag{A.1}
$$

where $\mu_s = E(\varepsilon^s)$, and the moments are calculated with the density of (2.5). These results can be derived by direct integration or, perhaps more conveniently, by means of the following orthogonal property of Hermite polynomials,

$$
\int_{-\infty}^{+\infty} \{H_m(x).H_n(x).f(x).\,\mathrm{d}x\} = \begin{cases} n! & , m = n \\ 0 & , \text{otherwise} \end{cases}
\tag{A.2}
$$

(see, for example, KENDALL and STUART [1977]).

The initial estimation of the a_s's required more elaboration : first, the variables w_s's were calculated; then, an OLS regression of ${u^2}_t$ on these w_s's variables provided the initial estimates for the a_s's values; the third step consisted of estimating a Normal density with conditional heteroskedasticity given as in (2.6). In this sequential way, initial estimates were obtained for all parameters, and the algorithms converged reasonably fast, and without further complications. The final step was the joint estimation of all the parameters involved, by a full Newton-Raphson algorithm, with analytical first and second order derivatives (the starting values being given by the procedure just described). Altogether, every estimated model required between 8 and 15 independent optimization processes, depending on the case. The required time, however, was not inadmissible (between 10 and 15 minutes with a P-133Mhz).

A visual comparison of the fit provided by the Normal and the ES densities is presented in Fig. 1. A more formal way of carrying out this comparison can be conducted by means of the following test,

$$
T.\Sigma_1{}^I[(f_i - p_i)^2 \,/\, p_i] \approx {X^2}_{I-1}
\tag{A.3}
$$

where $f_i = T_i / T$ is the observed frequency in a given interval, T_i being the number of observations falling in that interval, and T the total number of available observations in the sample. The probability, p_i, is the theoretical probability attached to that specific interval by the density being tested. The total number of intervals, I, covers the whole range of variations of the variable, so that intervals 1 and I, run to $-\infty$ and $+\infty$ respectively : in fact, this is rather convenient, since numerical problems make the denominator in the above expression equal to zero, at the far end of both tails. This test was implemented in the case analysed in Figure 1 ; I was set equal to 71 (69 equally spaced intervals between -3.4 and 3.4, plus the two tail intervals) ; the observations were *filtered* to eliminate the conditional heteroskedasticity, that is, u_t / σ_t (by means of the estimated a_s's, and using the expression of (2.6). If the probabilities p_i depend on a set of m parameters, and they are estimated from the raw data (not from the aggregated frequencies), and used to estimate the p_i's, then the test given above is asymptotically distributed as a *Chi-squared* variable under the null, with a number of degrees of freedom between $(I-1)$ and $(I-m-1)$ (see KENDALL and STUART [1977]). The value attained by this test was 4107 in the case of the Normal density, and 598 for the ES density ; assuming the most favourable situation, i.e., the number of degrees of freedom to be 70, the 95 % confidence value is 90 : therefore, the Normal specification is strongly rejected, and although the ES implies a huge reduction in the value of the test, it also leaves room for further improvements (this is not however, the focus of the present paper ; see MAULÉON and PEROTE [1997] for further discussion and results).

It should be pointed out, finally, that the series analysed were not the original rates of changes themselves but, rather, the residuals from an OLS regression of the following type,

$$\Delta \log(F_t) = b_1 . \Delta \log(F_{t-1}) + b_2 . \Delta \log(F_{t-2}) + b_3 . \Delta \log(J_{t-1}) + b_4 . \Delta \log(G_{t-1}) + b_5 . \Delta \log(US_{t-1}) + u_t \quad (A.4)$$

where F is the French stock market index, and J, G, US, similarly for Japan, Germany, and the United States (4 outright outliers were interpolated, as well ; similar equations were estimated for the remaining countries). The question arises, then, as to what extent this two steps procedure may affect the overall results. Nevertheless, consistency is not affected, the only effect being on the standard errors. In principle, the two steps estimation procedure is equivalent to imposing that certain parameters are constant ; therefore, this leads to an underestimation of the asymptotic standard errors ; however, joint estimation allows for an increase in the value of the likelihood, so that the combined finite sample effect is dubious. An example for the French data is provided in Table 5 : as can be seen from the results reported in that table, neither the coefficients, nor the *t-ratios* change too much. Since the joint estimation of all parameters was rather costly computationally, it was judged that, at the present stage, this would complicate excessively the research, without adding too much in exchange (the fit of the OLS regression was very low in all cases, anyway).

Table 3 –Estimation results for exchange rates

	France	Germany	Japan	USA
a_0	.529E-5 (12.1)	.743E-5 (14.0)	.666E-5 (7.7)	.767E-5 (5.5)
a_1	.9016 (19.7)	.829 (17.6)	.921 (20.6)	.871 (16.3)
a_2	-.902 (15.1)	-.819 (12.5)	-.838 (16.1)	-.871 (14.5)
a_3	.245 (7.4)	.137 (3.3)	.111 (6.8)	.074 (4.1)
a_4	-.038 (5.6)	-.011 (.9)	-.0087 (4.5)	-.00131 (.5)
a_5	.0031 (5.2)	.00046 (.4)	.000376 (4.0)	-.822E-4 (.6)
a_6	-.00011 (5.2)	-.00001 (.2)	-.784E-4 (3.9)	.33 (1.1)
a_7	.159E-5 (5.2)	.7E-7 (.07)	.612E-7 (3.9)	-.317 (1.3)
d_3	-.032 (5.0)	.037 (6.0)	-.034 (5.8)	-.016 (2.3)
d_4	.115 (12.5)	.127 (13.5)	.096 (12.4)	.071 (9.2)
d_6	.0183 (7.3)	.021 (8.5)	.015 (6.9)	.0081 (3.6)
d_8	.00229 (9.2)	.00245 (10.4)	.0019 (9.3)	.0014 (5.7)
T	5780 30/04/74 28/05/96	6240 28/06/72 28/05/96	6240 28/06/72 28/05/96	4890 31/08/77 28/05/96
LK	23700.6	25339.9	23553.4	18156.9
N	25	25	45	45

Table 4 – Estimation results for short term interest rates

	Japan	USA
a_0	.393E-4 (18.2)	.288E-4 (7.5)
a_1	.848 (25.0)	.845 (16.8)
a_2	-.643 (27.5)	-.736 (13.1)
a_3	.028 (10.1)	.0828 (.6)
a_4	-.000935 (9.9)	.007 (4.1)
a_5	.154E-4 (12.7)	-.511E-3 (6.1)
a_6	-.105E-6 (15.5)	.131E-4 (6.9)
a_7	.246E-9 (17.6)	-.113E-6 (7.1)
d_3	-.0375 (3.9)	.52E-3 (.1)
d_4	.118 (30.4)	.087 (20.4)
T	3451 07/04/82 28/06/95	3689 07/04/82 27/05/96
LK	11012.6	11756.6
N	85	45

Table 5 – Estimation results for the french stock market

	I	II
a_0	.0000345 (18.8)	.0000354 (17.9)
a_1	.834 (32.3)	.836 (26.3)
a_2	-.784 (29.1)	-.773 (27.2)
a_3	.092 (13.0)	.09 (12.2)
a_4	-.0039 (8.2)	-.0038 (7.7)
a_5	-.345E-5 (40.3)	-.350E-5 (38.4)
a_6	.366E-5 (5.6)	.356E-5 (5.4)
a_7	-.580E-7 (4.8)	-.564E-7 (4.6)
d_4	.126 (15.7)	.129 (15.8)
d_6	.0224 (10.0)	.0230 (10.2)
d_8	.00263 (11.9)	.00269 (12.0)
b_0	.000214 (1.6)	.000176 (1.8)
b_1	.0685 (5.4)	.115 (9.1)
b_2	-.0215 (1.8)	-.059 (4.9)
b_3	-.0397 (3.1)	-.0121 (1.2)
b_4	.103 (7.5)	.090 (8.2)
b_5	.307 (22.9)	.314 (30.1)

Note: I, the b's estimated independently.
II, joint estimation of all parameters.

A.2. The model for the *Long ARCH* conditional variance

It is convenient to start by reparameterizing the conditional variance model given in (2.6) as follows,

$$\begin{aligned}\sigma^2{}_t &= \phi_0 + \Sigma^N{}_1(\phi_s \,.\, u^2{}_{t-s}) \qquad , \phi_s \geq 0 \\ &= \phi_0 + \delta_0 \,.\, u^2{}_{t-s} + \Sigma^{N-1}{}_1[\delta_s \,.\, \Delta(u^2{}_{t-s})] \qquad \text{(A.5)}\end{aligned}$$

where the δ_s are solved next. First, for any constant, L, we write the identity,

$$\begin{aligned}\Sigma^N{}_1(\phi_s \,.\, L^s) &= \delta_0 \,.\, L + \Sigma^{N-1}{}_1[\delta_s \,.\, \Delta(L^s)] \\ &= \delta_0 \,.\, L + \delta_1 \,.\, (L - L^2) + \delta_2 \,.\, (L^2 - L^3) + \dots \\ &\quad + \delta_{N-2} \,.\, (L^{N-2} - L^{N-1}) + \delta_{N-1} \,.\, (L^{N-1} - L^N) \qquad \text{(A.6)}\end{aligned}$$

Equating powers on L^s, now, it is immediate that,

$$\begin{aligned}\delta_0 &= \Sigma^N{}_1\,(\phi_s) \\ \delta_1 &= \phi_1 - \delta_0 = -\Sigma^N{}_2\,(\phi_s) \\ \delta_2 &= \phi_2 + \delta_1 = -\Sigma^N{}_3\,(\phi_s) \\ \delta_3 &= \phi_3 + \delta_2 = -\Sigma^N{}_4\,(\phi_s) \\ &\dots, \\ \delta_i &= \phi_i + \delta_{i-1} = -\Sigma^N{}_{i+1} - \,(\phi_s) \qquad \text{(A.7)}\end{aligned}$$

and $\delta_{N-1} = -\phi_N$, which can be easily derived from the previous expressions. From these results we note that the following inequalities hold,

$$\delta_1 < \delta_2 < \delta_3, \dots < 0 \qquad \text{(A.8)}$$

which suggests that these coefficients lie on a smoothly decreasing (in absolute terms) function of s. One possible such a function is provided by the Almon specification for lag polynomials. Defining $a_0 = \phi_0$, $a_1 = \delta 0$, for the sake of notational clarity, the Almon function for the coefficients δ_s, is given as follows,

$$\delta_s = \Sigma^n{}_2(a_j \,.\, s^{j-2}), \qquad s \geq 1 \qquad \text{(A.9)}$$

from where,

$$\begin{aligned}\delta_1 &= a_2 + a_3 + \dots + a_n \\ \delta_2 &= a_2 + a_3 \,.\, 2 + a_4 \,.\, 2^2 + \dots + a_n \,.\, 2^{n-2} \qquad \text{(A.10)}\end{aligned}$$

This model can be substituted in (A.5), to yield,

$$\begin{aligned}\Sigma^{N-1}{}_1[\delta_s \,.\, \Delta\,(u^2{}_{t-s})] &= \Sigma^{N-1}{}_{s=1}[\,(\Sigma^n{}_{j=2}\,a_j \,.\, s^{j-2}) \,.\, \Delta(u^2{}_{t-s})\,] \\ &= \Sigma^n{}_{j=2}\,a_j \,.\, [\Sigma^{N-1}{}_{s=1}\,s^{j-2} \,.\, \Delta(u^2{}_{t-s})\,] \qquad \text{(A.11)}\end{aligned}$$

We can define, now, the variables w_j, as given by,

$$w_j = \Sigma^{N-1}{}_{s=1}[s^{j-2} \,.\, \Delta(u^2{}_{t-s})] \qquad , j = 2, \dots, n. \qquad \text{(A.12)}$$

and, in particular,

$$\begin{aligned}w_2 &= \Delta(u^2{}_{t-1}) + \Delta(u^2{}_{t-2}) + \dots + \Delta(u^2{}_{-N+1}) \\ w_3 &= \Delta(u^2{}_{t-1}) + \Delta(u^2{}_{t-2}) \,.\, 2 + \Delta(u^2{}_{t-3}) \,.\, 2^2 + \dots + \Delta(u^2{}_{-N+1}) \,.\, (N-1)^2 \\ &\dots, \qquad \text{(A.13)}\end{aligned}$$

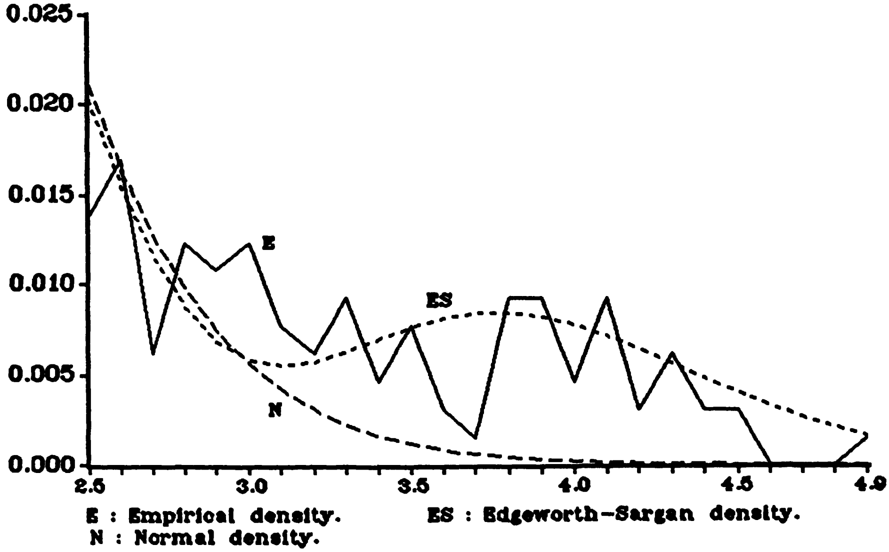
Finally, and substituting (A.12) in (A.11) and, in turn, (A.11) in (A.5), we get,

$$\sigma^2{}_t = a_0 + a_t \,.\, u^2{}_{t-1} + \Sigma^n{}_{j=2}\,(a_j \,.\, w_j) \qquad \text{(A.14)}$$

which is the expression given in the second line of (2.6) in the main text.

Fig. 3 – Tail densities

a) Upper tail



b) Lower tail

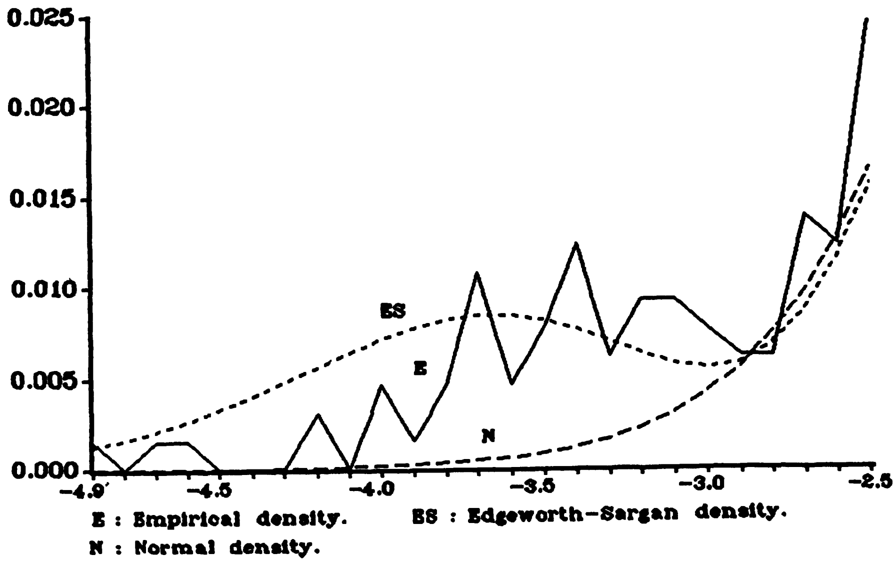
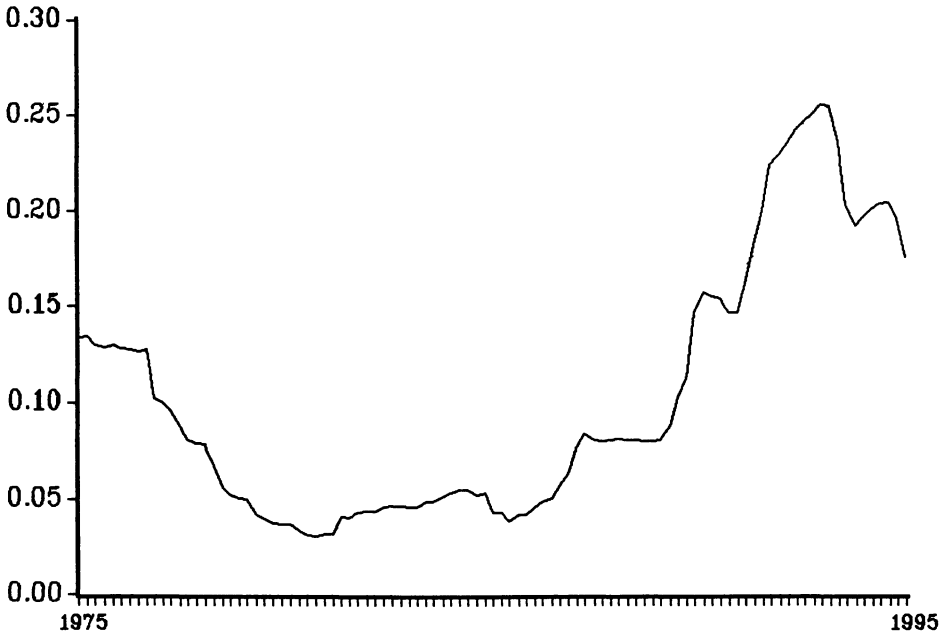


Fig. 4 – **Moving average of squared yields**

Japanese stock market



Moving average of the previous one thousand observations.
(squared rate of change of the stock market index)

References

BOURGOIN F. et PRIEUL D., (1997) "Estimation de diffusion à volatilité stochastique et application au taux d'intérêt à court terme français", *Journal de la Société de Statistique de Paris,* tome 138, n°4, pp. 43-66.

CAI J. (1994) "A Markow Model of Switching-Regime ARCH", *Journal of Business and Economic Statistics,* July 1994, vol. 12, N° 3, pp. 309-316.

DAVIDSON J. (1992) "A Central Limit Theorem for Globally Nonstationary Near-Epoch Dependent Functions of Mixing Processes", *Econometric Theory,* 8, pp. 313-329.

DING Z., GRANGER C. et ENGLE R. (1993) "A Long Memory Model of Stock Market Returns and a New Model", *Journal of Empirical Finance,* 1, pp. 107-131.

ENGLE R. et BOLLERSLEV T. (1986) "Modelling the Persistence of Conditional Variances", *Econometric Reviews,* 5, pp. 1-50.

GALLANT R. et NYCHKA D. (1987) "Seminonparametric Maximum Likelihood Estimation", *Econometrica,* 55, pp. 363-390.

GOODHART C. et O'HARA M. (1996) "High Frequency Data in Financial Markets : Issues and Applications", *Journal of Empirical Finance,* (forthcoming).

HAMILTON J. (1994) "Time Series Analysis", *Princeton.*

HANSEN B. (1995) "Regression with Nonstationary Volatility", *Econometrica,* Vol. 63, N° 5, pp. 1113-1132.

HARVEY A. et ROBINSON P. (1991) "Efficient Estimation of Nonstationary Time Series Regression", *Journal of Time Series Analysis,* 9, pp. 201-214.

KENDALL M. et STUART A. (1977) *The Advanced Theory of Statistics,* Griffith & Co., London, 4th ed.

LASTRAPES W. (1989) "Exchange Rate Volatility and U.S. Monetary Policy : an ARCH Application", *Journal of Money, Credit and Banking,* Vol. 21, N° 1, pp. 66-77.

LAMOUREUX CH. et LASTRAPES W. (1990) "Persistence in Variance, Structural Change, and the GARCH Model", *Journal of Business and Economic Statistics,* Vol. 8, N° 2, pp. 225-234.

MAULEÓN I. (1983) *Approximations to the Finite Sample Distribution of Econometric Chi. Squared Criteria,* Ph.D. dissertation, *London School of Economics.*

MAULEÓN I. et PEROTE J. (1997) "Testing Densities with Financial Data : an Empirical Comparison of the Edgeworth-Sargan Density to the Student's t", Proceedings of the *Forecasting Financial Markets conference, Imperial College,* London.

NELSON D. (1991) "Conditional Heteroskedasticity in Asset Returns : a New Approach", *Econometrica,* 59, pp. 347-370.

PAGAN A. et SCHWERT G. (1990a) "Alternative Models for Conditional Stock Volatility", *Journal of Econometrics,* 45, pp. 267-290.

PAGAN A. et SCHWERT G. (1990b) "Testing for Covariance Stationarity in Stock Market Data", *Economics Letters,* 33, pp. 165-170.

PERRON P. (1989) "The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis", *Econometrica,* 57, pp. 1361-1401.

PHILLIPS P. et LORETAN M. (1994) "Testing the Covariance Stationarity of Heavy-Tailed Time Series", *Journal of Empirical Finance,* 1, pp. 211-248.

PHILLIPS P. et LORETAN M. (1995) "On the Theory of Testing Covariance Stationarity under Moment Condition Failure", *in Advances in Econometric and Quantitative Economic*, Eds. G. Maddala, P. Phillips, T. Srinivasan, Blackwell, Oxford, pp. 198-233.

SARGAN J.D. (1976) "Econometric Estimators and the Edgeworth Approximation", *Econometrica,* 44, pp. 421-448.

SOWELL F. (1992) "Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models", *Journal of Econometrics,* 8.

WOOLRIDGE J. et WHITE H. (1988) "Some Invariant Principles and Central Limit Theorems for Dependent Heterogeneous Processes", *Econometric Theory,* 4, pp. 210-230.