JANUARY 1990
REVISED NOVEMBER 1990

N° 9109

QUALITATIVE THRESHOLD ARCH MODELS

C. GOURIEROUX*, A. MONFORT**

(*) CREST-ENSAE et CEPREMAP

(**) INSEE - Département de la Recherche

QUALITATIVE THRESHOLD ARCH MODELS

ABSTRACT

In this paper we consider a class of dynamic models in which both the conditional mean and the conditional variance are endogenous stepwise functions. We first consider the probabilistic properties of these models : stationarity conditions, leptokurtic effect, linear representation, optimal prediction ; in this first part most results are based on Markov chains theory. Then we derive statistical properties of this class of models : pseudo-maximum likelihood estimators, conditional homoscedasticity tests, tests of weak or strong white noise, CAPM test, factors determination, ARCH-M effects. We also discuss the introduction of exogenous variables and the case of multiple lags. Finally, an application to the Paris Stock Index is proposed.

MODELES ARCH A SEUILS QUALITATIFS

RESUME

Dans cet article nous considérons une classe de modèles dynamiques dans lesquels la moyenne et la variance conditionnelle sont des fonctions endogènes constantes par morceaux. On considère d'abord les propriétés probabilistes de ces modèles : conditions de stationarité, effet leptokurtique, représentation linéaire, prédiction optimale ; dans cette première partie la plupart des résultats sont fondés sur la théorie des chaînes de Markov. Ensuite on établit les propriétés statistiques de cette classe de modèles : estimateurs du pseudo-maximum de vraisemblance, tests d'homoscédasticité conditionnelle, tests de bruit blanc faible et fort, test du CAPM, détermination de facteurs, effet ARCH-M. On discute également l'introduction de variables exogènes et le cas de retards multiples. Finalement on propose une application à l'indice CAC.

Keywords : ARCH models - Financial Assets - Heteroscedasticity

Mots Clés : Modèles ARCH - Actifs financiers - Hétéroscédasticité.

J.E.L. CLASSIFICATION SYSTEM : 210.

I - INTRODUCTION

The time series literature on the univariate or multivariate ARMA models provided new approaches of the dynamic econometric modelling. In particular, the VAR models are widely used either as an alternative of the structural models (Sims 1980) or as a framework in which a sequence of tests can be performed in order to select a structural model [Hendry-Mizon (1990), Monfort-Rabemananjara (1990)]. However, this literature is basically interested in the conditional mean (given the past), which is assumed to be linear, and makes the strong assumption that the conditional variance is fixed. This drawback has been stressed by Engle (1982) and his work initiated a stream of papers on ARCH on GARCH models (Bollerslev (1986)). In this literature the conditional variance is very often specified as a linear function of the squared values of past innovations, even if non parametric approaches have also been proposed (see Gregory (1989), Engle-Gonzalves-Rivera (1989)).

The present paper deals with several issues. First we explore the possible trade-off between the flexibility of the conditional variance specification in terms of a given past value and the number of relevant lags. Secondly we adopt a symmetric teatment of the conditional mean and the conditional variance in order to discuss the possible cross effects of mispecifications. Thirdly, like in the VAR approach, we propose a general framework and statistical methods allowing for the tests of various restrictions ; however contrary the VAR approach, these restrictions may concern both the conditional mean and the condition variance. Finally, since we do not wish to make parametric distributional assumptions, like conditional normality, we propose to use pseudo-likelihood techniques [Gourieroux-Monfort Trognon (1984)].

Since we are interested in simple flexible parameterizations of the conditional mean and the conditional variance, there are two natural candidates for the classes of functional forms : the piecewise constant functions and the piecewise linear functions. In this paper we consider piecewise constant functions which have the advantage of being also available in the multivariate case ; piecewise linear functions are used in Zakoian (1990). More precisely the basic model considered in this paper is, in the case of one lag :

$$Y_t = \sum_{j=1}^{J} \alpha_j \mathbb{1}_{A_j}(Y_{t-1}) + \sum_{j=1}^{J} \beta_j \mathbb{1}_{A_j}(Y_{t-1}) u_t$$

where Y_t is the multivariate series of interest, $\{A_j, j=1,\dots,J\}$ is a partition of the set of values of Y, α_j is an unknown vector, β_j is an unknown symmetric positive definite matrix, $\mathbb{1}_{A_j}$ is the characteristic function of A_j and (u_t) is a strong white noise. This kind of model can be seen as a generalization of the threshold models for the conditional mean [Tong-Lim (1980), Tong (1983), Chan-Petrucelli-Woolford (1985), Saikkonen-Luukkonen (1986), Melard-Roy (1987)].

In section 2, we derive the stochastic properties of the process Y and of the innovation process $\sum_{j=1}^{J}\beta_j \mathbb{1}_{A_j}(Y_{t-1})u_t$.This study is based on a preliminary study of the underlying qualitative process $Z_t = (\mathbb{1}_{A_1}(Y_t),\dots,\mathbb{1}_{A_J}(Y_t))'$, which is a regular Markov chain under weak conditions. We obtain the expressions of the mean and of the autocovariance function of Y, we examine the leptokurtic effect induced by the conditional heteroscedasticity. We also prove that the process Y has a linear ARMA(J-1,J-1) representation as well as (in the univariate case) Y_t^2 and the squared error process v_t^2, whith $v_t = \sum_{j=1}^{J}\beta_j \mathbb{1}_{A_j}(Y_{t-1})$.These properties allow to discuss the consequences of various specification errors.

In section 3, we give the expression and the asymptotic properties of the pseudo-maximum likelihood estimators of the parameters α_j and β_j. Then we describe the tests procedures of a number of hypotheses : hypotheses on the partition, homoscedasticity hypothesis, weak or strong white noise hypothesis, ARCH-M hypothesis, CAPM hypothesis, factors and efficiency hypotheses. All these hypotheses are easily tested by using the asymptotic least squares theory [Gourieroux-Monfort-Trognon (1985), Gourieroux-Monfort-Renault (1988), Gourieroux-Monfort (1989-b)]

In section 4, we consider several generalizations in particular the introduction of exogenous variables, and the case of several lags. In section 5 we propose an application on the Paris Stock Index (indice CAC).

II - DEFINITIONS AND PROBABILISTIC PROPERTIES

II.1 Definition of the model QTARCH(1)

The process of interest $\{Y_t,\ t\in\mathbb{Z}\}$ is n-dimensional and satisfies

(1) $$Y_t = \sum_{j=1}^{J} \alpha_j \mathbb{1}_{A_j}(Y_{t-1}) + \sum_{j=1}^{J} \beta_j \mathbb{1}_{A_j}(Y_{t-1}) u_t$$

where $\{A_j,\ j\in J\}$ is a partition of $\mathbb{R}^n$, α_j, $j=1,\ldots,J$ are n dimensional vectors, β_j, $j=1,\ldots,J$ are positive definite matrices and $\{u_t,\ t\in\mathbb{Z}\}$ a sequence of i.i.d. unobservable random vectors whose mean and covariance matrix are respectively zero and idendity. This model where only one lag appears is called a QTARCH(1) model.

It is also assumed that the probability distribution of any u_t, denoted by Q, is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^n$ and that its p.d.f, denoted by g, is strictly positive. An important particular case is the normal case (i.e. the case where $u_t \sim N(0,I_n)$ but this normality assumption will not be made except when explicitly mentioned.

From $\{Y_t,\ t\in\mathbb{Z}\}$ it is possible to define the J-multivariate process $Z_t = (Z_{1t},\ldots,Z_{Jt})'$ with :

(2) $$Z_{jt} = \mathbb{1}_{A_j}(Y_t) \ .$$

This process Z can be considered as a qualitative process with J possible states. Moreover Y_t can be rewritten in the following way :

(3) $$Y_t = \alpha' Z_{t-1} + \beta' (Z_{t-1} \otimes I_n)\, u_t \ ,$$

where $\alpha' = (\alpha_1,\ldots,\alpha_J)$ and $\beta' = (\beta_1,\ldots,\beta_J)$. Note that, if e is J-vector whose entries are all equal to 1, we have $e'Z_t = 1$.

It is also worth stressing that the α_j's are not necessarily different and that the same is true for the β_j's ; this implies that the relevant partitions for the mean and the variance can always been assumed to be identical since, if they are different, we obtain a model of type (1) by considering the intersection of these partitions.

II.2 Stationarity

It is easy to prove the following proposition showing that it is sufficient to study the stationarity of the qualitative process Z.

Proposition 1

$\{Y_t,\ t\in\mathbb{Z}\}$ is strictly stationary if and only if $\{Z_t,\ t\in\mathbb{Z}\}$ is strictly stationary.

The qualitative process Z is an homogeneous Markov chain of order one whose transition matrix is denoted by P.

The (j,k) entry of this matrix is :

$$p_{jk} = \Pr(Z_{jt} = 1/Z_{k,t-1} = 1)$$
$$= \Pr(Y_t \in A_j/Z_{k,t-1} = 1)$$
$$= \Pr(\alpha_k+\beta_k u_t \in A_j) \ ,$$

(4) $$p_{jk} = Q[\beta_k^{-1}(A_j-\alpha_j)] \ .$$

The assumptions made above imply that all the transition probabilities p_{jk} are strictly positive and, therefore, the transtion matrix P is completely regular (see Gantmacher (1966) chapter 13). From Perron's theorem we deduce that 1 is a single eigenvalue of P and that a corresponding eigenvector π can be chosen with all its entries strictly positive and with $\pi' e = 1$. π is the invariant probability of the Markov chain.

The other eigenvalues λ_j, $j=1,\dots,J-1$ have a modulus strictly smaller than one and P can be written using the spectral decomposition (see appendix 1) :

(5) $$P = \pi e' + \sum_{j=1}^{J-1} \lambda_j a_j b_j' \ ,$$

with (using the notation $a_J=\pi$, $b_J=e$).

$$b_j' a_j = 1 \qquad j = 1,\dots,J \ ,$$
$$b_k' a_j = 0 \qquad \forall j \neq k, \quad k,j = 1,\dots,J \ ,$$

and therefore, the matrices $C_j = a_j b_j'$ are idempotent.

The marginal probability p_t of Z_t is given by :

(6) $$p_t = P^t p_0 \ .$$

If p_0 is equal to π, the same is true for any p_t ; moreover, since the chain is completely regular, p_t converges to π when t goes to infinity for any p_0. This is also a consequence of the equality :

(7) $$P^t = \pi e' + \sum_{j=1}^{J-1} \lambda_j^t a_j b_j' \ ,$$

since $$P^t p_0 = \pi + \sum_{j=1}^{J-1} \lambda_j^t a_j b_j' p_0 \text{ converges to } \pi \ .$$

Using proposition 1 we immediately get :

Corollary 2

If $p_0 = \pi, \{Y_t,\ t \in \mathbb{Z}\}$ is strictly stationary ; otherwise this process is asymptotically strictly stationary.

The invariant p.d.f. of Y_t is :

(8) $$\sum_{j=1}^{J} \frac{\pi_j}{\det \beta_j} g(\beta_j^{-1}(Y_t-\alpha_j)) \quad .$$

III.3 Unconditional moments.

Under the stationarity assumption we have :

$$E\,Y_t = E(\alpha' Z_{t-1} + \beta'(Z_{t-1} \otimes I_n)u_t)$$
$$= E(\alpha' Z_{t-1}) \ ,$$

(9) $$E\,Y_t = \alpha'\pi = \sum_{j=1}^{J} \pi_j \alpha_j .$$

The unconditional covariance matrix of Y_t is :

$$V\,Y_t = V\,E(Y_t/Z_{t-1}) + E\,V(Y_t/Z_{t-1})$$
$$= V(\alpha' Z_{t-1}) + E(\beta'\ (Z_{t-1}Z_{t-1}' \otimes I_n)\beta) \ ,$$

(10) $$V\,Y_t = \alpha'(\text{diag}\ \pi - \pi\pi')\alpha + \beta'(\ \text{diag}\ \pi \otimes I_n)\beta .$$

This matrix can also be written :

$$V\,Y_t = \sum_{j=1}^{J} \pi_j(\alpha_j\alpha_j' + \beta_j^2) - \left(\sum_{j=1}^{J}\pi_j\alpha_j\right)\left(\sum_{j=1}^{J}\pi_j\alpha_j\right)' \ ,$$

and it reduces to : $\sum_{j=1}^{J} \pi_j(\alpha_j\alpha_j'+\beta_j^2)$, if the process Y is zero-mean.

It can also be shown that the autocovariance function of Y_t is :

(11) $$\gamma_Y(h) = \text{cov}(Y_t,Y_{t-h}) = \sum_{j=1}^{J-1} \lambda_j^{h-1}\alpha' a_j b_j' C, \ h \geq 1 \text{ with } C = E(Z_t Y_t')$$

Let us now focus on the error term of model (1) :

(12) $$v_t = \sum_{j=1}^{J} \beta_j \mathbb{1}_{A_j}(Y_{t-1})u_t \ ,$$

and let us consider the univariate case.

If we denote $\beta = (\beta_1,\ldots,\beta_J)'$, we see that the process v is a second order (weak) white noise, whose variance is :

$$E\,v_t^2 = E(\beta' Z_{t-1})^2$$
$$= \beta'(\text{diag}\ \pi)\beta$$
$$= \sum_{j=1}^{J}\beta_j^2\pi_j \ .$$

The fourth order moment of v_t is :

$$E\,v_t^4 = E[(\beta' Z_{t-1})^4 u_t^4]$$
$$= E\,u_t^4 E(\beta' Z_{t-1})^4$$
$$= \mu_4 \sum_{j=1}^{J}\beta_j^4\pi_j \ ,$$

where μ_4 is the fourth order moment of u_t, or its kurtosis since $Eu_t^2=1$, and $Eu_t=0$.

The kurtosis of v_t is measured by :

$$k = \frac{E\ v_t^4}{(E\ v_t^2)^2} = \mu_4 \frac{\sum_{j=1}^{J} \beta_j^4 \pi_j}{\left(\sum_{j=1}^{J} \beta_j^2 \pi_j\right)^2} ,$$

(13)
$$k = \mu_4 \left[1 + \frac{V_\pi \beta_j^2}{(E_\pi \beta_j^2)^2}\right] ,$$

where E_π and V_π denote the empirical mean and variance with respect to π. It is clear that k is greater than μ_4. In particular if u_t is normal, v_t is leptokurtic. From (13) it appears that the kurtosis of v_t increases when that of u_t increases but also when the relative variability of the conditional variances increases. So the kurtosis depends on a natural measure of the heterogeneity in the conditional variances β_j.

It is also interesting to study the properties of the v_t^2 process, whose mean is $\sum_{j=1}^{J} \beta_j^2 \pi_j$. After some algebra the autocovariance function of v_t^2 can be shown to be :

(14)
$$\text{cov}(v_t^2, v_{t-h}^2) = (\beta_1^2, \ldots, \beta_J^2) \sum_{j=1}^{J-1} a_j b_j' \lambda_j^{h-1} d.$$

with $d = E[Z_t (\beta_1^2, \ldots, \beta_J^2) Z_{t-1} u_t^2]$

II.4. <u>Optimal prediction</u>

The optimal prediction of $Y_{t+h} (h \geq 1)$ given $(\underline{Y}_t = \{Y_t, Y_{t-1}, \ldots\}$ is the conditional expectation :

(15)
$$\begin{aligned} E(Y_{t+h}/\underline{Y}_t) &= E(\alpha' Z_{t+h-1} + \beta'(Z_{t+h-1} \otimes I_n) u_{t+h}/\underline{Y}_t) \\ &= E(\alpha' Z_{t+h-1}/\underline{Y}_t) \\ &= \alpha' E(Z_{t+h-1}/\underline{Z}_t) \\ &= \alpha' P^{h-1} Z_t , \end{aligned}$$

This optimal prediction is clearly a nonlinear function of $\underline{Y}_t$ and it can also be written, using (7) :

$$E(Y_{t+h}/\underline{Y}_t) = \alpha' \pi e' Z_t + \sum_{j=1}^{J-1} \lambda_j^{h-1} \alpha' a_j b_j' Z_t ,$$

(16)
$$E(Y_{t+h}/\underline{Y}_t) = \alpha' \pi + \sum_{j=1}^{J-1} \lambda_j^{h-1} \alpha' a_j b_j' Z_t ,$$

When h goes to infinity $E(Y_{t+h}/\underline{Y}_t)$ converges to $\alpha' \pi = E\ Y_t$.
Note that these optimal forecasts depend not only on the mean parameters α, but also on the variance parameters β_j.

Similarly the conditional covariance matrix can be written :

$$V(Y_{t+h}/\underline{Y}_t) = V[E(Y_{t+h}/\underline{Y}_{t+h-1})/\underline{Y}_t] + E[V(Y_{t+h}/\underline{Y}_{t+h-1})/\underline{Y}_t]$$
$$= V[\alpha' Z_{t+h-1}/\underline{Y}_t] + E[\beta'(Z_{t+h-1}Z'_{t+h-1} \otimes I_n)\beta/\underline{Y}_t] .$$

The conditional mean and covariance matrix of Z_{t+h-1} given $\underline{Y}_t$ are respectively $P^{h-1}Z_t = \mu_{t,h-1}$ (say) and $\text{diag}(\mu_{t,h-1})-\mu_{t,h-1}\mu'_{t,h-1}$. So we have :

$$(17) \qquad V(Y_{t+h}/\underline{Y}_t) = \alpha'(\text{diag } \mu_{t,h-1}-\mu_{t,h-1}\mu'_{t,h-1})\alpha + \beta'(\text{diag } \mu_{t,h-1} \otimes I_n)\beta.$$

When h goes to infinity $\mu_{t,h-1} = P^{h-1}Z_t$ converges to $\pi\ e'Z_t = \pi$, and therefore $V(Y_{t+h}/\underline{Y}_t)$ converges to

$$\alpha'(\text{diag } \pi-\pi\pi')\alpha+\beta'(\text{diag } \pi\otimes I_n)\beta=VY_t.$$

II.5. <u>Linear representations and autocovariance function</u>

Let us consider the following zero-mean strictly and second order stationary process :

$$W_t = \prod_{j=1}^{J-1} (1-\lambda_j L)(Y_t-EY_t) ,$$

where L is the lag operator and the λ_j's the eigenvalues of the transition matrix different from one.

All the components of Y appearing in W_t have an index τ satisfying $t-J+1\leq\tau\leq t$. The conditional expectation $E(W_t/\underline{Y}_{t-h})$, $h\geq J$, can be computed from (12) and we get :

$$E(W_t/\underline{Y}_{t-h}) = \prod_{j=1}^{J-1} (1-\lambda_j L) \sum_{i=1}^{J-1} \lambda_i^{h-1}\alpha' a_i b'_i Z_{t-h} ,$$

where the lag operator L operates only on the power of the λ_i's, i.e.:

$$L^k\left(\sum_{i=1}^{J-1} \lambda_i^{h-1}\alpha' a_i b'_i Z_{t-h}\right) = \sum_{i=1}^{J-1} \lambda_i^{h-k-1}\alpha' a_i b'_i Z_{t-h} .$$

From the previous expression we conclude that $E(W_t/\underline{Y}_{t-h}) = 0$, for any $h\geq J$ or, equivalently, $E(W_t/\underline{W}_{t-h}) = 0$, for any $h\geq J$. This implies, in particular, that the process W_t has a moving average representation of order J-1 and, consequently, that Y_t has an ARMA(J-1, J-1) representation with a scalar autoregressive operator.

Moreover $\prod_{j=1}^{J-1}(1-\lambda_j L)$ is clearly equal to $\dfrac{\det(I-PL)}{1-L}$.

Formula (14) shows that in the univariate case v_t^2 has also an ARMA(J-1, J-1) representation with the same autoregressive operator. It can also be shown that the same is true for Y_t^2.

Proposition 3

Under the stationarity assumption, $\{Y_t, t\in\mathbb{Z}\}$ has a linear ARMA(J-1, J-1) representation of the form :

$$\frac{\det(I-PL)}{1-L}(Y_t - E\, Y_t) = \epsilon_t + \sum_{j=1}^{J-1} \theta_j \epsilon_{t-j} .$$

In the univariate case the same is true for $\{v_t^2, t\in\mathbb{Z}\}$ and $\{Y_t^2, t\in\mathbb{Z}\}$.

The previous proposition has several consequences in terms of specification. First, if $\{Y_t, t\in\mathbb{Z}\}$ is a QTARCH(1) and if a linear specification is chosen, the number of relevant lags in a AR or MA specification may be large (infinite in theory), particularly if some eigen values λ_j have a modulus close to 1. This clearly shows that there exists a tradeoff between the number of relevant lags and the degree of non linearity. Similarly, if v_t is the error process of a QTARCH and if it is specified as a GARCH(p,q) process and identified through the ARMA[max(p,q),p] representations of v_t^2, the values chosen for p or q may be high. In particular if a $|\lambda_j|$ is near 1, an ARCH(q) with a large q may be selected although, by definition, only one lag is relevant for the conditional variance; an IGARCH model may also be selected. Finally the previous proposition stresses the need of a simultaneous modelling of the non linearities appearing in the conditional mean and the conditional variance, in order to avoid cross effects of misspecifications ; for instance if the true process $\{Y_t, t\in\mathbb{Z}\}$ is a conditionnally homoscedastic process ($\beta_1 = \beta_2 = \ldots = \beta_J$) and if it is specified as a ARCH model, identified through the ARMA representation of Y_t^2, a strong ARCH effect could be found.

Example 1

Let us consider the univariate QTARCH(1) process defined by :

(18) $$Y_t = -\alpha_1 \mathbb{1}_{\mathbb{R}^+}(Y_{t-1}) - \alpha_2 \mathbb{1}_{\mathbb{R}^-}(Y_{t-1}) + [\beta_1 \mathbb{1}_{\mathbb{R}^+}(Y_{t-1}) + \beta_2 \mathbb{1}_{\mathbb{R}^-}(Y_{t-1})]u_t$$

$\beta_1, \beta_2 > 0$, $u_t \sim IID(0,1)$.

The transition matrix P is :

$$P = \begin{bmatrix} P[Y_t>0/Y_{t-1}>0] & P[Y_t>0/Y_{t-1}<0] \\ P[Y_t<0/Y_{t-1}>0] & P[Y_t<0/Y_{t-1}<0] \end{bmatrix}$$
$$= \begin{bmatrix} P[-\alpha_1+\beta_1 u_t>0] & P[-\alpha_2+\beta_2 u_t>0] \\ P[-\alpha_1+\beta_1 u_t<0] & P[-\alpha_2+\beta_2 u_t<0] \end{bmatrix}$$
$$= \begin{bmatrix} 1-G(\gamma_1) & 1-G(\gamma_2) \\ G(\gamma_1) & G(\gamma_2) \end{bmatrix},$$

where G is the cumulative distribution function of u_t and

$$\gamma_1 = \frac{\alpha_1}{\beta_1}, \quad \gamma_2 = \frac{\alpha_2}{\beta_2} .$$

The eigenvalues of P are 1 and $G(\gamma_2)-G(\gamma_1)$, and the invariant probability (π_1,π_2) is given by :

$$\begin{cases} \pi_1 = [1-G(\gamma_2)]/[1-G(\gamma_2)+G(\gamma_1)] , \\ \pi_2 = G(\gamma_1)/[1-G(\gamma_2)+G(\gamma_1)]. \end{cases}$$

The mean of Y_t is :

$$E\, Y_t = \frac{-\alpha_1(1-G(\gamma_2))-\alpha_2 G(\gamma_1)}{1-G(\gamma_2)+G(\gamma_1)} .$$

The linear representation of Y is an ARMA(1,1) of the following form :

$$\begin{aligned} Y_t-[G(\gamma_2)-G(\gamma_1)]Y_{t-1} &= [1-G(\gamma_2)+G(\gamma_1)]E\, Y_t+\varepsilon_t+\theta\varepsilon_{t-1} \\ &= -\alpha_1[1-G(\gamma_2)]-\alpha_2 G(\gamma_1)+\varepsilon_t+\theta\varepsilon_{t-1} . \end{aligned}$$

The parameter θ and the variance of ε_t can be determined using the expressions of $V\, Y_t$ and $Cov(Y_t,Y_{t-1})$. If $\beta_1 = \beta_2$, Y_t is a weak white noise and a cancellation appears between the AR and the MA polynomials.

The linear representation of v_t^2 and Y_t^2 are similar.

Example 2

Let us consider the following QTARCH(1) model

$$Y_t = \sum_{j=1}^{10} \alpha_j \mathbb{1}_{A_j}(Y_{t-1}) + \left[\sum_{j=1}^{10} \beta_j \mathbb{1}_{A_j}(Y_{t-1})\right] u_t \tag{19}$$

where u_t is IIN(0,1) and the partitioning $\{A_j,\ j=1,10\}$ is defined by the intervals whose boundary points are -4, -3, -2, -1, 0, 1, 2, 3, 4. The conditional standard errors are assumed to be equal to one in the central intervals [-1,0] and [0,1], i.e. $\beta_5 = \beta_6 = 1$; the other values of the β_j's are parameterized by $j=\beta_1$ and $k=\beta_{10}$, the other values being deduced by linear functions from β_1 to β_5 and from β_6 to β_{10}. Let us first assume that all the α_j are equal to zero. Figure 1 shows the largest modulus of the eigenvalues of the transition matrix (once the eigenvalue 1 has been excluded) as a function of j and k. It is seen that this modulus may be high, particularly for large values of j and k. This suggests that, although the model is markovian of order one, an ARCH(q) specification, in which the conditional variance is a linear function of the squared past values, may necessitate a large number of lags. Figure 2 is similar to the previous one when $\alpha_1=\alpha_2=\ldots=\alpha_5=-1.5$ and $\alpha_6=\alpha_7=\ldots=\alpha_{10}=1.5$ and similar conclusions can be drawn in this case for small values of j and k (the horizontal axes are decreasing) ; note however that cancellations in the AR and MA polynomials may appear (in particular if j = k = 1).

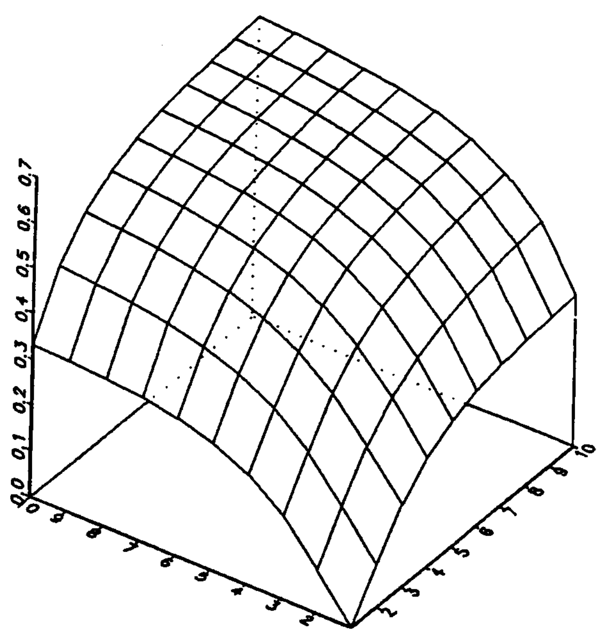

Figure 1

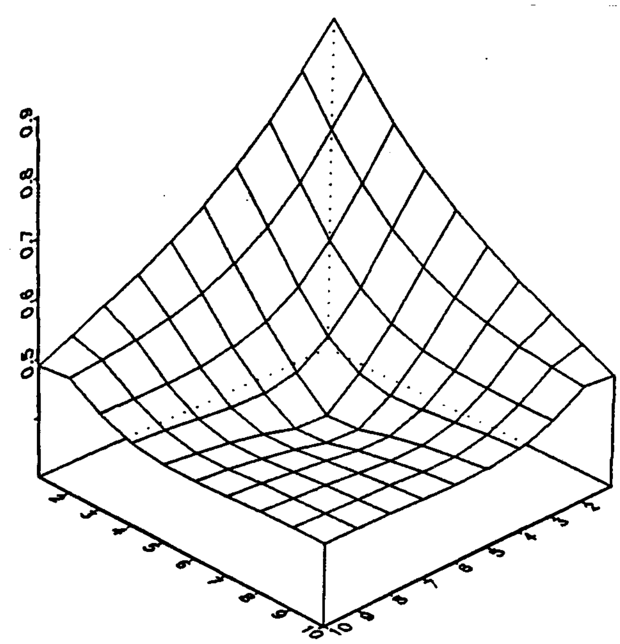

Figure 2

III. STATISTICAL PROPERTIES

III.1. The pseudo-maximum likelihood estimators

Let us assume that we have observed $Y_0,\ldots,Y_T$ and that the process u is IID(0,1) ,not necessarily normal. The parameters $\alpha_j, \beta_j, j=1,\ldots,J$ can be estimated by the pseudo-maximum likelihood method based on the normal distribution. The P.M.L. estimators are the solutions of :

(19) $$\underset{\alpha,\beta}{\text{Max}}\ L_T = \sum_{j=1}^{J} \sum_{t\in B_j} \left\{ -\frac{n}{2} \text{Log } 2\pi - \text{Log det } \beta_j - \frac{1}{2}(Y_t-\alpha_j)'\beta_j^{-2}(Y_t-\alpha_j) \right\}.$$

where B_j is the set defined by $\{t: 1\le t\le T,\ Y_{t-1}\in A_j\}$.
If T_j is the cardinal of B_j, we get :

(20) $$L_T = \sum_{j=1}^{J} \left\{ -\frac{nT_j}{2} \text{Log } 2\pi - T_j \text{ Log det } \beta_j - \frac{1}{2} \sum_{t\in B_j} (Y_t-\alpha_j)'\beta_j^{-2}(Y_t-\alpha_j) \right\}$$

The maximisation of L_T can be made separately with respect to the (α_j,β_j), $j=1,\ldots,J$ and we get the least squares estimators :

(21) $$\begin{cases} \hat{\alpha}_j = \bar{Y}_j = \dfrac{1}{T_j} \sum_{t\in B_j} Y_t , \\ \hat{\beta}_j^2 = \dfrac{1}{T_j} \sum_{t\in B_j} (Y_t-\bar{Y}_j)(Y_t-\bar{Y}_j)' . \end{cases}$$

These estimates are empirical mean and and variance, but the sets B_j, on which these empirical moments are computed, are endogenous. In the case of rate of returns, a usual way of determining expected returns and volatilities consists in averaging on H consecutive observations, i.e. in determining :

$$\begin{cases} \tilde{\alpha}_t = \dfrac{1}{H} \sum_{\tau=t}^{t+H} Y_\tau , \\ \tilde{\beta}_t^2 = \dfrac{1}{H} \sum_{\tau=t}^{t+H} (Y_\tau-\tilde{\alpha}_t)(Y_\tau-\tilde{\alpha}_t)' . \end{cases}$$

It is clear that the previous estimators $\hat{\beta}_j^2$, can be interpreted in terms of conditional (or instantaneous) volatilities, whereas the usual estimators $\tilde{\beta}_t^2$ are marginal (or historical) volatilities.

The same kind of comments applies to functions of the parameters α,β. For instance, we may consider a bidimensional model, where the first series is the cum-dividend return on a given security and the second one is the market return. The "beta" of this security is often evaluated through a regression of Y_{1t} on Y_{2t}, i.e. is estimated by the empirical regression coefficient associated with $\text{Cov}(Y_{1t},Y_{2t})/V(Y_{2t})$; it is an "historical" beta. In fact a more useful definition of the beta is the regression coefficient conditional to the past $\text{Cov}(Y_{1t}, Y_{2t}/\underline{Y}_{t-1})/V(Y_{2t}/\underline{Y}_{t-1})$. In the QTARCH model the beta is β_{12j}/β_{22j}, if $t\in B_j$, where β_{12j} and β_{22j} are (1,2) and (2,2) entries of β_j. The betas vary with time in an endogenous way.

III.2. The asymptotic properties of the estimators

These properties are summarized in the following proposition proved in appendix 2.

Proposition 4 The P.M.L. estimators of the α_j's and β_j^2's are asymptotically normal; they are also asymptotically independent if $Eu_t^3 = 0$. The asymptotic covariances matrices of $\sqrt{T}(\hat{\alpha}_j-\alpha_j)$ and $\sqrt{T}[\text{vec } \hat{\beta}_j^2 - \text{vec } \beta_j^2]$ are such that :

$$V_{as}[\sqrt{T}(\hat{\alpha}_j-\alpha_j)] = \beta_j^2/\pi_j \ ,$$

and :

$$Cov_{as}[\sqrt{T}(\hat{B}_{klj}-B_{klj}), \sqrt{T}(\hat{B}_{k^*l^*j}-B_{k^*l^*j})] = \frac{1}{\pi_j} C_{klk^*l^*j} \ ,$$

where $C_{klk^*l^*j} = Cov(\beta_j^{k'} u_t u_t' \beta_j^{l}, \beta_j^{k^{*\prime}} u_t u_t' \beta_j^{l^*})$ and β_j^k is the column of matrix β_j.

In the univariate case the asymptotic variance of $\sqrt{T}(\hat{\beta}_j^2-\beta_j^2)$ reduces to :

$$\frac{1}{\pi_j} \beta_j^4(\mu_4-1) \ ,$$

where μ_4 is the kurtosis of u_t

In practice the asymptotic variances can be easily estimated from their empirical counterparts. For instance since

$C_{klk^*l^*j} = Cov(v_{kt}v_{lt}, v_{k^*t}v_{l^*t}/Y_{t-1} \in A_j)$, where v_{kt} is the k^{th} component of the innovation, a consistent estimator is :

$$\hat{C}_{klk^*l^*j} = \frac{1}{T_j} \sum_{t\in B_j} \hat{v}_{kt}\hat{v}_{lt}\hat{v}_{k^*t}\hat{v}_{l^*t} - \left(\frac{1}{T_j} \sum_{t\in B_j} \hat{v}_{kt}\hat{v}_{lt}\right) \times \left(\frac{1}{T_j} \sum_{t\in B_j} \hat{v}_{k^*t}\hat{v}_{l^*t}\right)$$

where $\hat{v}_{kt}$ is the residual associated with v_{kt}.

III.3. Estimation and test when the mean and variance partitions are different

When the partitions for the mean, $\{A_k^M, \ k=1,\dots,K_M\}$, and the variance, $\{A_k^V, \ k=1,\dots,K_V\}$ are different, it is possible to use the previous model with the intersection of these partitions denoted by $\{A_j, \ j=1\dots J\}$ and to estimate the α_j's and the β_j's by the previous pseudo-maximum likelihood method. However this method does not take into acount the constraints on the α_j's and the β_j's induced by the equalities :

$$(22) \qquad \begin{cases} A_k^M = \bigcup_{j\in J_M^k} A_j \ , & k = 1,\dots,K_M \ , \\ A_k^V = \bigcup_{j\in J_V^k} A_j \ , & k = 1,\dots,K_V \ , \end{cases}$$

that is to say the constraints :

$$(23)\qquad \begin{cases} \alpha_j = \alpha_{j'}\ , \ \forall_{j,j'} \in J^k_M\ , \ k = 1,\dots,K_M\ , \\ \beta^2_j = \beta^2_{j'}\ , \ \forall_{j',j} \in J^k_V\ , \ k = 1,\dots,K_V\ . \end{cases}$$

The constrained estimators can be obtained by the asymptotic least squares method. In the univariate case, we get :

$$(24)\qquad \bar{\alpha}_k = \sum_{j\in J^k_M} \frac{\hat{\alpha}_j T_j}{\hat{\beta}^2_j} \Big/ \sum_{j\in J^k_M} \frac{T_j}{\hat{\beta}^2_j}\ , \ k = 1,\dots,K_M$$

$$(25)\qquad \bar{\beta}^2_k = \sum_{j\in J^k_V} \frac{T_j}{\hat{\beta}^2_j} \Big/ \sum_{j\in J^k_V} \frac{T_j}{\hat{\beta}^4_j}\ .$$

An asymptotically optimal test statistic of these constraints is the value at the optimum of the O.L.S. objective function, i.e. :

$$(26)\qquad S_1 = \sum_{k=1}^{K_M} \sum_{j\in J^k_M} \frac{T_j(\hat{\alpha}_j-\bar{\alpha}_k)^2}{\hat{\beta}^2_j} + \sum_{k=1}^{K_V} \sum_{j\in J^k_V} \frac{T_j(\hat{\beta}^2_j-\bar{\beta}^2_k)^2}{\hat{\beta}^4_j(\hat{\mu}_4-1)}$$

Under the null hypothesis, given in (23), S_1 is asymptotically distributed as a chi-square with $2J-K_M-K_V$ degrees of freedom.
In summary :

<u>Proposition 5</u>
The partitions $\{A^M_k,\ k=1,\dots,K_M\}$ and $\{A^V_k,\ k=1,\dots,K_V\}$ are accepted at the asymptotic level ϵ, if $S_1 \leq \chi^2_{1-\epsilon}(2J-K_M-K_V)$.

Also note that, under H_0, the asymptotic distributions of $\sqrt{T}(\bar{\alpha}_k-\tilde{\alpha}_k)$ (where $\tilde{\alpha}_k$ is the common values of the α_j for $j\in J^k_M$) and $\sqrt{T}(\bar{\beta}^2_k-\tilde{\beta}^2_k)$ (where $\tilde{\beta}^2_k$ is the common value of the β^2_j for $j\in J^k_V$, are given by :

$$(27)\qquad \sqrt{T}(\bar{\alpha}_k-\tilde{\alpha}_k) \xrightarrow[T\to\infty]{L} N\left[0,\ \left(\sum_{j\in J^k_M} \frac{\pi_j}{\beta^2_j}\right)^{-1}\right]$$

and

$$(28)\qquad \sqrt{T}(\bar{\beta}^2_k-\tilde{\beta}^2_k) \xrightarrow[T\to\infty]{L} \left[0,(\mu_4-1)\ \left(\sum_{j\in J^k_V} \frac{\pi_j}{\beta^4_j}\right)^{-1}\right]\ .$$

All these variables are asymptotically independent.
It is also worth noting that since the distribution of u_t is generally non normal, the ALS estimator $\bar{\beta}^2_k$ is asymptotically more efficient than the constrained pseudo-maximum likelihood estimator [see Gourieroux-Monfort (1989-a, chapter X)].

III.3. <u>Test of conditional homoscedasticity</u> (in the univariate case).

Let us first assume that the partitions for the mean and the variance are identical. In this case the conditional homoscedasticity is characterized by :

(29) $$H_0 : \beta_1^2 = \ldots = \beta_J^2 .$$

From the previous subsection, it is clear that, under H_0, the estimator of the common value of the β_j^2, denoted by $\bar{\beta}^2$, is :

(30) $$\bar{\beta}^2 = \sum_{j=1}^{J} \frac{T_j}{\hat{\beta}_j^2} \Big/ \sum_{j=1}^{J} \frac{T_j}{\hat{\beta}_j^4} .$$

The test statistic is :

(31) $$S_2 = \sum_{j=1}^{J} \frac{T_j(\hat{\beta}_j^2 - \bar{\beta}^2)^2}{\hat{\beta}_j^4(\hat{\mu}_4 - 1)} ,$$

whose asymptotic distribution under H_0 is $\chi^2(J-1)$.
If the partitions are different the homoscedasticity assumption is :

(32) $$H_0 : \tilde{\beta}_1^2 = \ldots = \tilde{\beta}_{K_v}^2$$

and the estimator of their common value is :

(33) $$\bar{\beta}_*^2 = \sum_{k=1}^{K_v} \bar{\beta}_k^2 \sum_{j \in J_v^k} \frac{T_j}{\hat{\beta}_j^4} \Big/ \sum_{j=1}^{J} \frac{T_j}{\hat{\beta}_j^4}$$

and from (25) :

$$\bar{\beta}_*^2 = \sum_{j=1}^{J} \frac{T_j}{\hat{\beta}_j^2} \Big/ \sum_{j=1}^{J} \frac{T_j}{\hat{\beta}_j^4}$$

which, as expected, is the same estimator as the one obtained directly from the unconstrained estimator β_j^2 associated with the intersection of the two partitions.

The test statistic is :

(34) $$S_3 = \sum_{k=1}^{K_v} \frac{(\bar{\beta}_k^2 - \bar{\beta}^2)^2}{(\hat{\mu}_4 - 1)} \sum_{j \in J_v^k} \frac{T_j}{\hat{\beta}_j^4}$$

whose asymptotic distribution, under H_0, is $\chi^2(K_v-1)$.
In summary :

Proposition 6

If the mean and variance partitions are the same, the conditional homoscedasticity assumption is rejected at the level ε if $S_2 \geq \chi^2_{1-\varepsilon}(J-1)$, where S_2 is given in (31). If the partitions are different this assumption is rejected if $S_3 \geq \chi^2_{1-\varepsilon}(K_V-1)$, where S_3 is given in (34).

III.4. Test of weak and strong white noise

Let us first consider the case where the mean and variance partitions are identical. Moreover, in order to simplify the notations we consider the univariate case.

If $\alpha_1 = \alpha_2 = \ldots = \alpha_J$, the process Y_t is a weak (or second order) white noise whose mean is the common value of the α_j's. Using the same approach as in the previous subsection an A.L.S. based test statistic for this null hypothesis is :

$$(35) \qquad S_4 = \sum_{j=1}^{J} \frac{T_j(\hat{\alpha}_j - \bar{\alpha})^2}{\hat{\beta}_j^2},$$

where

$$\bar{\alpha} = \sum_{j=1}^{J} \frac{\hat{\alpha}_j T_j}{\hat{\beta}_j^2} \Big/ \sum_{j=1}^{J} \frac{T_j}{\hat{\beta}_j^2},$$

whose asymptotic distribution under the null is $\chi^2(J-1)$. If the common value of the α_j's is zero, Y_t is a zero-mean weak white noise and this hypothesis can be tested from the statistic :

$$(36) \qquad S_5 = \sum_{j=1}^{J} \frac{T_j \hat{\alpha}_j^2}{\hat{\beta}_j^2}$$

whose distribution under the null is $\chi^2(J)$.

If the mean and variance partitions are different, S_4 and S_5 are replaced, respectively, by :

$$(37) \qquad S_6 = \sum_{k=1}^{K_M} (\bar{\alpha}_k - \bar{\alpha})^2 \sum_{j \in J_M^k} \frac{T_j}{\hat{\beta}_j^2},$$

and :

$$(38) \qquad S_7 = \sum_{k=1}^{K_M} \bar{\alpha}_k^2 \sum_{j \in J_M^k} \frac{T_j}{\hat{\beta}_j^2},$$

whose asymptotic distributions under the null are respectively $\chi^2(K_M-1)$ and $\chi^2(K_M)$.

The strong white noise property is characterized by the equality of the conditional means and the equality of the conditional variances; moreover in the zero-mean strong white noise case we assume that the common value of the conditional means is zero.

When the partitions are the same, the strong white noise hypothesis and the zero-mean strong white noise hypothesis are respectively tested from the statistics S_2+S_4 and S_2+S_5, whose asymptotic distributions under the null are, respectively, $\chi^2(2J-2)$ and $\chi^2(2J-1)$.
When the partitions are different, the relevant statistics are S_3+S_6 and S_3+S_7, whose asymptotic distributions under the null are respectively $\chi^2(K_V+K_M-2)$ and $\chi^2(K_V+K_M-1)$. In summary, we have the following proposition.

Proposition 7

At the asymptotic level ε, we have the following critical regions :

- if the partitions are identical,
  - weak white noise hypothesis : $S_4 \geq \chi^2_{1-\varepsilon}(J-1)$
  - zero-mean weak white noise : $S_5 \geq \chi^2_{1-\varepsilon}(J)$
  - strong white noise : $S_2+S_4 \geq \chi^2_{1-\varepsilon}(2J-2)$
  - zero-mean strong white : $S_2+S_5 \geq \chi^2_{1-\varepsilon}(2J-1)$
- if the partitions are different
  - weak white noise : $S_6 \geq \chi^2_{1-\varepsilon}(K_M-1)$
  - zero-mean weak white noise : $S_7 \geq \chi^2_{1-\varepsilon}(K_M)$
  - strong white noise : $S_5+S_6 \geq \chi^2_{1-\varepsilon}(K_V+K_M-2)$
  - zero-mean strong white noise hypothesis : $S_3+S_7 \geq \chi^2_{1-\varepsilon}(K_V+K_M-1)$

III.5. Tests on ARCH-M effects

Engle-Lilien-Robbins (1987) introduced the notion of ARCH-M model in which the conditional variance or the conditional standard error appears in the conditional mean.

In our univariate model this kind of condition implies that the partitions for the mean and the variance are the same and that :

$$\exists \gamma : \alpha_j = \gamma \beta_j^2 \quad j=1,\dots,J \text{ (variance case)}$$

or

$$\exists \gamma : \alpha_j = \gamma \beta_j \quad j=1,\dots,J \text{ (standard error case)}$$

Let us first consider the variance case. The hypothesis $\{\exists \gamma : \alpha_j = \gamma \beta_j^2\}$ is in a mixed form and is easily tested (see Gourieroux-Monfort (1989-b)), in the following way.

In a first stage we compute the OLS estimator of γ, from the artificial regression :

$$\hat{\alpha}_j = \gamma\hat{\beta}_j^2 + w_j \quad j=1,\dots,J ,$$

and we get :

$$\tilde{\gamma} = \frac{\sum_{j=1}^{J} \hat{\alpha}_j \hat{\beta}_j^2}{\sum_{j=1}^{J} \hat{\beta}_j^4} .$$

Then we compute the asymptotic variances of $\sqrt{T}(\hat{\alpha}_j-\gamma\beta_j^2)$ under the null, i.e., $\frac{\beta_j^2}{\pi_j}(1+\gamma^2\beta_j^2(\mu_4-1))$, which is estimated by $\frac{T\beta_j^2}{T_j}[1+\tilde{\gamma}^2\beta_j^2(\hat{\mu}_4-1)]$. The test statistic is T times the optimal value of the objective function in the previous artifial regression when the GLS is applied with the variances given above. We get the statistic :

(39) $$S_8 = \sum_{j=1}^{J} (\hat{\alpha}_j-\hat{\gamma}\hat{\beta}_j^2)^2 \frac{T_j}{\hat{\beta}_j^2[1+\tilde{\gamma}^2\hat{\beta}_j^2(\hat{\mu}_4-1)]},$$

with

$$\hat{\gamma} = \sum_{j=1}^{J} \frac{\hat{\alpha}_j T_j}{1+\tilde{\gamma}^2\hat{\beta}_j^2(\hat{\mu}_4-1)} \Big/ \sum_{j=1}^{J} \frac{\hat{\beta}_j^2 T_j}{1+\tilde{\gamma}^2\hat{\beta}_j^2(\hat{\mu}_4-1)},$$

whose asymptotic distribution under the null is $\chi^2(J-1)$.

In the standard error case we use the result :

$$\sqrt{T}(\hat{\beta}_j-\beta_j) \xrightarrow[T\to\infty]{L} N\left[0, \frac{\beta_j^2}{4\pi_j}(\mu_4-1)\right],$$

and the statistic becomes :

$$S_9 = \sum_{j=1}^{J} (\hat{\alpha}_j-\gamma^*\hat{\beta}_j)^2 \frac{T_j}{\hat{\beta}_j^2+\tilde{\tilde{\gamma}}^2 \frac{\beta_j^2}{4}(\hat{\mu}_4-1)},$$

(40) $$S_9 = 4 \sum_{j=1}^{J} (\hat{\alpha}_j-\gamma^*\hat{\beta}_j)^2 \frac{T_j}{\hat{\beta}_j^2(4+\tilde{\tilde{\gamma}}^2\hat{\mu}_4-\tilde{\tilde{\gamma}}^2)},$$

with :

$$\gamma^* = \sum_{j=1}^{J} \frac{\hat{\alpha}_j T_j}{\hat{\beta}_j} \Big/ \sum_{j=1}^{J} T_j$$

$$= \frac{1}{T} \sum_{j=1}^{T} \frac{\hat{\alpha}_j}{\hat{\beta}_j} T_j,$$

and $$\tilde{\tilde{\gamma}} = \frac{\sum_{j=1}^{J} \hat{\alpha}_j \hat{\beta}_j}{\sum_{j=1}^{J} \hat{\beta}_j^2}$$

The asymptotic distribution of S_9 under the null is $\chi^2(J-1)$.

Proposition 8

The critical regions at the asymptotic level ϵ are :

$S_8 \geq \chi^2_{1-\epsilon}(J-1)$ for the ARCH-M effect in variance,

$S_9 \geq \chi^2_{1-\epsilon}(J-1)$ for the ARCH-M effect in standard error.

III.6. Tests of the CAPM

Let us now consider tests based on financial theories.

The Capital Asset Pricing Model is based on the assumptions that individual portfolios are determined in an optimal way and that there is a clearing condition assuring that the market portfolio is a convex combination of individuals' optimal portfolio. It follows immediately that the market portfolio is on the portfolio frontier [see Huang-Litzenberger (1988)]. This condition implies some restrictions on dynamic models describing the excess rates of return with respect to that of a riskless asset.

a) First, if $(Y_{1t},\ldots,Y_{nt})'$ are such net rates of return, if $\lambda=(\lambda_1,\ldots,\lambda_n)'$ is a vector whose entries are the supplies in the different assets assumed to be fixed (as mentioned by Engle-Ng-Rothschild (1989) it is a strong assumption), we have under the CAPM :

$$E(Y_t/\underline{Y}_{t-1}) = a\, V(Y_t/\underline{Y}_{t-1})\lambda ,$$

where a is a constant coefficient measuring the risk aversion. Distinguishing the different regimes of our model, we get :

$$\exists\ a : \alpha_j = a\, \beta_j^2 \lambda \quad j=1,\ldots,J.$$

This kind of hypothesis might be tested along the following lines if the quantities λ are available and exogenous.

i) In the first stage, we regress by O.L.S. $\hat{\alpha}_j$ on $\hat{\beta}_j^2\lambda$ for the different regimes, which gives an estimate of the measure of the risk aversion $\tilde{a}$, which is consistent under the null.

ii) Then we estimate the asymptotic covariance matrix :

$$\Gamma(a,\lambda,\alpha_j,\beta_j^2) = V_{as}[\sqrt{T}(\hat{\alpha}_j - a\hat{\beta}_j^2\lambda)] ,$$

by using property 4 and replacing the unknown asymptotic variances of $\sqrt{T}(\hat{\alpha}_j-\alpha_j)$, $\sqrt{T}(\text{vec } \hat{\beta}_j^2 - \text{vec } \beta_j^2)$ by their estimates and a by $\tilde{a}$; let us denote by $\hat{\Gamma}_j$ the matrix thus obtained.

iii) In the second stage, we regress $\hat{\alpha}_j$ on $\hat{\beta}_j^2\lambda$ by G.L.S., using $\hat{\Gamma}_j$ as covariance matrices, and we get a better estimate $\hat{a}$ of coefficient a.

iv) The test is based on the statistic :

$$S_{1e} = T \sum_{j=1}^{J} (\hat{\alpha}_j - \hat{a}\hat{\beta}_j^2\lambda)' \hat{\Gamma}_j^{-1} (\hat{\alpha}_j - \hat{a}\hat{\beta}_j^2\lambda) \tag{41}$$

Using the results in Szroeter (1983) Gourieroux-Monfort-Renault (1988) and Gourieroux-Monfort (1989-b) we get :

Proposition 9

The CAPM hypothesis is rejected, at the asymptotic level ϵ, if $S_{1e} \geq \chi^2_{1-\epsilon}(nJ-1)$.

Note that if λ is unknown it is be possible to implement the same kind of test since the constraints become $\exists\lambda^*: \alpha_j = \beta_j^2\lambda^*$ and the degrees of freedom of the statistic obtained is $n(J-1)$.

From a descriptive point of view, the residual plots, i.e. the values of $\hat{\alpha}_j - \hat{a}\hat{\beta}_j^2\lambda$, may be informative. They may allow to detect some regions for which the CAPM is not satisfied, i.e. the j values for wich $\hat{\alpha}_j - \hat{a}\hat{\beta}_j^2\lambda$ is "far" from zero.

b) The CAPM is often tested from some of its consequences . For instance it is known that the CAPM also implies some restrictions in which λ does not directly appear. If the net rates of return of the assets and of the market portfolio are both available, we may write a joint model on $(Y_{ot}, Y_{1t} \dots Y_{nt})$, where o is the index for the market. In such a case the CAPM implies :

$$E\left[\begin{pmatrix} Y_{1t} \\ \vdots \\ Y_{nt} \end{pmatrix} / \underline{Y}_{t-1}\right] = \text{Cov}\left[\begin{pmatrix} Y_{1t} \\ \vdots \\ Y_{nt} \end{pmatrix}, Y_{ot}/\underline{Y}_{t-1}\right] V(Y_{ot}/\underline{Y}_{t-1})^{-1} E(Y_{ot}/\underline{Y}_{t-1})$$

If we introduce the block decompositions :

$$\alpha_j = \begin{pmatrix}\alpha_{0j}\\ \tilde{\alpha}_j\end{pmatrix}, \quad \beta_j^2 = \begin{pmatrix} B_{00j} & \tilde{B}_{0j} \\ \tilde{B}_{j0} & \tilde{B}_j \end{pmatrix},$$

the condition gives the following implicit restrictions on the parameters

(42) $$\tilde{\alpha}_j = \tilde{B}_{j0}(B_{00j})^{-1}\alpha_{0j} .$$

These restrictions can be tested in the usual way by Wald's procedure either regime by regime, i.e. separetely for the different j's, or globally for all the j's.

III.6. <u>Factors determination and efficiency</u>

It is interesting, in multivariate financial time series models, to look for directions, i.e. linear combinations (or portfolios) of the initial series with specific properties. The QTARCH models may be useful for an empirical determination of such directions or factors. Let us consider for example the determination of conditionally homoscedastic directions [Diebold-Nerlove (1986), (1989), Engle (1987) Engle-Ng-Rothschild (1989), Nerlove-Diebold-Van Beek-Cheung (1988)]. Conditional homoscedasticity exists for a given portfolio associated with the weights μ iff :

$\mu'\beta_j^2\mu$ does not depend or j.

This hypothesis may also be written under a mixed form :

$$\exists\mu\in\mathbb{R}^n, \mu\neq 0 \ \exists\nu\in\mathbb{R}^+ : \mu'\beta_j^2\mu=\nu, \ \forall j$$

or

(43) $$\exists\mu\in\mathbb{R}^n : \ \mu'\beta_j^2\mu=1 \ \forall j \ \text{(if the } \beta_j^2 \text{ are invertible)}$$

This hypothesis can be tested using a generalized Wald test [see Szroeter (1983), Gourieroux-Monfort (1989-b)], if J is greater than n.

In a first stage, we first determine a consistent estimator $\tilde{\mu}$ of μ by minimizing :

$$\sum_{j=1}^{J} (\mu'\hat{\beta}_j^2\mu-1)^2 .$$

Then we determine the asymptotic variance of $\sqrt{T}(\mu'\hat{\beta}_j^2\mu-1)$ which is a fonction of μ and of the asymptotic covariance matrix of vec $\hat{\beta}_j^2$. This asymptotic variance can be consistently estimated by replacing $\tilde{\mu}$ by μ and $V_{as}\sqrt{T}(\text{vec } \hat{\beta}_j^2 - \text{vec } \beta_j^2)$ by the estimate based in property 5. We denote $\hat{\gamma}_j$ this estimated variance.

Finally the generalized Wald statistic is defined by :

$$S_{11} = T \operatorname*{Min}_{\mu} \sum_{j=1}^{J} \frac{1}{\hat{\gamma}_j} [\mu' \hat{\beta}_j^2 \mu - 1]^2 \ .$$

Using the results established in Szroeter and Gourieroux-Monfort, we get the following result.

Proposition 10: If we know that there exists at most one vector $\mu \in \mathbb{R}^n$ such that $\mu' \beta_j^2 \mu = 1$ and if $J>n$, the generalized Wald statistic S_{11} is asymptotically distributed under the null hypothesis as a chi-square distribution with J-n degrees of freedom.

A given set of K portfolios is defined by the (Kxn) matrix of its weights, denoted by B. It is readily seen that, if the process Y_t of the net rates of returns, is a QTARCH process the efficiency condition of the set of portfolios B is :

(44) $\exists \ \nu_j \in \mathbb{R}^K \ : \ (\beta_j^2)^{-1} \alpha_j = B' \nu_j \ , \ \forall j$.

This hypothesis has a bilinear mixed form can be easily tested by using a method similar to that proposed above for the CAPM [see Gourieroux-Monfort-Renault (1988)] ; this method leads to a statistic asymptotically distributed as a $\chi^2[(n-K)J]$ under H_0.

If we want to test that there exists an efficient set of K(<J) portfolios (this set being no longer given) the assumption can be written in the same way except that B is unknown ; moreover, for identifiability reasons, B depends on (n-K)xK independent parameters and therefore the test procedure for bilinear mixed assumptions leads to a statistic whose asymptotic distribution under H_0 is $\chi^2[(n-K)(J-K)]$. An equivalent way of writing this assumption is :

(45) $\exists C \ : \ C(\beta_j^2)^{-1} \alpha_j = 0 \quad \forall j$

where C is a (n-K)xn matrix depending on (n-K)K parameters.

IV. EXTENSIONS

IV.1. Exogenous variables

Let us now assume that an exogenous vector appears in the right hand side of equation (1). More precisely let us assume that the model is univariate, for notational simplicity, and defined by :

$$(46) \qquad Y_t = \sum_{j=1}^{J} (\alpha_j + x_t a_j)\, \mathbb{1}_{A_j}(Y_{t-1}) + \sum_{j=1}^{J} \beta_j \mathbb{1}_{A_j}(Y_{t-1}) u_t,$$

where x_t is row vector of size L and a_j a column vector of L parameters. We assume that the process $\{x_t\}$ is independent from the process $\{u_t\}$.

If $\{x_t\}$ is stationary, $\{Y_t\}$ is stationary as soon as $\{Z_t\}$ is stationary. The qualitative process $\{Z_t\}$ is a Markov chain whose transition matrix P* is defined by :

$$\begin{aligned} p^*_{jk} &= \Pr[Y_t \in A_j / Z_{k,t-1} = 1] \\ &= \Pr[\alpha_k + x_t a_k \in A_j] \\ &= \underset{x}{E}\, Q\left(\frac{A_j - \alpha_k - x a_k}{\beta_k}\right) \end{aligned}$$

and we get the same kind of results as in section II if P is replaced by P*.

The pseudo-log likelihood function is :

$$(47) \qquad \begin{aligned} L^*_T &= \sum_{j=1}^{J} \sum_{t\in B_j} \left[-\frac{1}{2} \operatorname{Log}(2\pi\beta_j^2) - \frac{1}{2\beta_j^2} (y_t - \alpha_j - x_t a_j)^2 \right] \\ &= \sum_{t=1}^{T} l_t \end{aligned}$$

with

$$l_t = \sum_{j=1}^{J} \mathbb{1}_{A_j}(y_{t-1}) \left[-\frac{1}{2} \operatorname{Log}(2\pi\beta_j^2) - \frac{1}{2\beta_j^2} (y_t - \alpha_j - x_t a_j)^2 \right]$$

It is always possible to center the x_t vectors within each class of index B_j ; in other words, if X_j denotes the $T_j \times L$ matrix whose rows are x_t, $t\in B_j$, we assume that the sum of the elements of any column is zero. In this case, it is easily seen that the pseudo-maximum likelihood estimators of the α_j's are the same as in III.1, and those of the a_j's are :

$$(48) \qquad \hat{a}_j = (X_j' X_j)^{-1} X_j' Y_{(j)}$$

where $Y_{(j)}$ is the vector whose components are Y_t, $t\in B_j$ (in the same order as the rows of X_j). The pseudo M.L. estimator of β_j^2 becomes :

$$(49) \qquad \hat{\beta}_j^2 = \frac{1}{T_j} \sum_{t\in B_j} (Y_t - \bar{Y}_j - x_t \hat{a}_j)^2$$

It can be seen that :

$$\frac{1}{T}\sum_{t=1}^{T} E\left[\frac{\partial^2 l_t}{\partial\alpha_j \partial a_j} / x_t\right] = 0 \ , \ \frac{1}{T}\sum_{t=1}^{T} E\left[\frac{\partial^2 l_t}{\partial\beta_j^2 \partial a_j} / x_t\right] = 0,$$

and, if $Eu_t^3 = 0$: $\frac{1}{T}\sum_{t=1}^{T} E\left[\frac{\partial l_t}{\partial\alpha_j} \frac{\partial l_t}{\partial a_j} / x_t\right] = 0, \ \frac{1}{T}\sum_{t=1}^{T} E\left[\frac{\partial l_t}{\partial\beta_j^2} \frac{\partial l_t}{\partial a_j} / x_t\right] = 0.$

These results imply that the variables $\sqrt{T}(\hat{\alpha}_j-\alpha)$, $\sqrt{T}(\beta_j^2-\beta_j^2)$, $\sqrt{T}(\hat{a}_j-a_j)$ $j=1,\dots,J$ are asymptotically independent.
Moreover :

(50) $$\sqrt{T}(\hat{\alpha}_j-\alpha) \xrightarrow[T\to\infty]{L} N\left(0, \frac{\beta_j^2}{\pi_j}\right) ,$$

(51) $$\sqrt{T}(\hat{a}_j-a_j) \xrightarrow[T\to\infty]{L} N\left[0, (Exx')^{-1} \frac{\beta_j^2}{\pi_j}\right] ,$$

(52) $$\sqrt{T}(\hat{\beta}_j-\beta_j^2) \xrightarrow[T\to\infty]{L} N\left[0, \frac{\beta_j^4}{\pi_j} (\mu_4-1)\right] .$$

The various tests on the α_j's, β_j's proposed in the previous section remain valid.

An additional test would be a test of linearity with respect to x_t, i.e. the test of the null hypothesis : $a_1=a_2\dots=a_J$. The common value a of the a_j's can be estimated from the A.L.S. model :

$$\begin{cases} \hat{a}_1 = a + w_1 \\ \quad \vdots \\ \hat{a}_J = a + w_J \end{cases}$$

with $$V(w_j) = \frac{T\hat{\beta}_j^2}{T_j} (X'X)^{-1}, \text{ where } X' = (X_1',\dots,X_J')$$

We obtain :

$$\hat{a} = \left[\frac{1}{T} X'X \sum_{j=1}^{J} \frac{T_j}{\hat{\beta}_j^2}\right]^{-1} \frac{1}{T} X'X \sum_{j=1}^{J} \frac{T_j}{\hat{\beta}_j^2} \hat{a}_j$$

(53) $$\hat{a} = \sum_{j=1}^{J} \frac{T_j}{\hat{\beta}_j^2} \hat{a}_j \Big/ \sum_{j=1}^{J} \frac{T_j}{\hat{\beta}_j^2}$$

The asymptotic covariance matrix of $\sqrt{T}(\hat{a}-a)$ is $\left[\sum_{j=1}^{J} \frac{\pi_j}{\hat{\beta}_j^2} Ex'x\right]^{-1}$,

which is estimated by $\left[\frac{1}{T^2} \sum_{j=1}^{J} \frac{T_j}{\hat{\beta}_j^2} X'X\right]^{-1}$.

The test statistic is :

$$(54) \qquad S_{12} = \frac{1}{T}\sum_{j=1}^{J} \frac{T_j}{\hat{\beta}_j^2} (\hat{a}_j-\hat{a})'X'X(\hat{a}_j-\hat{a})$$

Proposition 11

The asymptotic critical region, for testing the non linearity in x at the asymptotic level ϵ, is $S_{12} \geq \chi^2_{1-\epsilon}[L(J-1)]$.

IV.2. Multiple lags

The statistical methods proposed above can be extended to multiple lags. The more general model in this case is :

$$(55) \qquad Y_t = \sum_{j=1}^{J} \alpha_j \mathbb{1}_{A_j}(Y_{t-1},\dots,Y_{t-p})u_t + \sum_{j=1}^{J} \beta_j \mathbb{1}_{A_j}(Y_{t-1},\dots,Y_{t-p})u_t$$

where $\{A_j,\ j=1,\dots,J\}$ is a partition of $\mathbb{R}^p$.

The main problem which is likely to arise in this case is the large number of parameters. In order to reduce the number of parameters, it is possible to assume first that the partition $\{A_j, j=1,\dots,J\}$ is the product of a partition in $\mathbb{R}$ $\{A_i^*, i=1,\dots,I\}$; in this case model can be written :

$$(56) \qquad Y_t = \sum_{i_1=1}^{I}\dots\sum_{i_p=1}^{I} \alpha_{i_1,\dots,i_p} \mathbb{1}_{A^*_{i_1}\times\dots\times A^*_{i_p}}(Y_{t-1},\dots,Y_{t-p}) + \left[\sum_{i_1=1}^{I}\dots\sum_{i_p=1}^{I} \beta_{i_1,\dots,i_p} \mathbb{1}_{A^*_{i_1}\times\dots\times A^*_{i_p}}(Y_{t-1},\dots,Y_{t-p})\right]u_t$$

In this kind of specification it is possible to adopt an approach which is similar to the analysis of variance. In particular a significant reduction of the number of parameters will be obtained by assuming an additive model, or a model without time interactions. For instance, in the univariate case such a model can be written :

$$(57) \qquad Y_t = \alpha_0 + \sum_{i=1}^{I-1}\sum_{j=1}^{p} \alpha_{ij} \mathbb{1}_{A_i^*}(Y_{t-j}) + \left[\beta_0 + \sum_{i=1}^{I-1}\sum_{j=1}^{p} \beta_{ij} \mathbb{1}_{A_i^*}(Y_{t-j})\right]u_t$$

with the positivity constraints, $\beta_0>0$ and $\beta_0 + \beta_{i_1,1}+\dots+\beta_{i_p,p}>0$ for any $(i_1,\dots,i_p)$. Within this framework, it would be also possible to test a more restrictive model, defined in the same spirit as the GARCH model and called generalized QTARCH or G-QTARCH :

$$(58) \qquad \begin{cases} Y_t = m_t + \sigma_t u_t \\ m_t = a_0 m_{t-1} + \sum_{i=1}^{I} a_i \mathbb{1}_{A_i^*}(Y_{t-1}) \\ \sigma_t^2 = \delta_0 \sigma_{t-1}^2 + \sum_{i=1}^{I} \delta_i \mathbb{1}_{A_i^*}(Y_{t-1}),\ \delta_i>0 \end{cases}$$

V - AN APPLICATION

Let us now illustrate the previous results by investigating the conditional variance of the daily relative change of the Paris stock index (indice CAC), from January 86 to April 90. As a first insight in the data, let us consider a QTARCH(2) model where the space (Y_{t-1}, Y_{t-2}) is partitioned into 36 sets obtained from the product of the univariate partition whose bounds are -0.8%, -0.4%, 0%, 0.4%, 0.8%. This model can be written :

$$(59) \qquad Y_t = \left[\sum_{i=1}^{6} \sum_{j=1}^{6} \beta_{ij} \mathbb{1}_{A_i}(Y_{t-1}) \mathbb{1}_{A_j}(Y_{t-2})\right] u_t$$

The (pseudo) maximum likelihood estimators of the β_{ij} are all significant, using a one-sided 5% ratio test based either on the ML standard errors or on the PML standard errors. The estimation of the fourth moment of u_t based on this model is 4.1, suggesting a leptokurtic effect and justifying the use of a PML approach. A few conditional means computed with the same partition are marginally significant but, in the sequel, we concentrate on the conditional variances and the conditional means are taken equal to zero.

The estimates $\hat{\beta}_{ij}$ of the β_{ij} are shown in figure 3 (note that the coordinates of the horizontal axes are the centers of the intervals A_i and A_j in decreasing order). This figure seems to show that the conditional standard errors β_{ij} are increasing functions of the absolute values of the interval centers corresponding to Y_{t-1} and Y_{t-2}; however it seems that the responses are not symmetrical for the negative and for the positive values of Y_{t-1} and Y_{t-2}. In particular, figure 4 shows the values of the $\hat{\beta}_{ii}$ $(i=1,\dots,6)$ and indicates that the conditional standard errors are larger for the negative values of Y_{t-1} or Y_{t-2} than for the positive values.

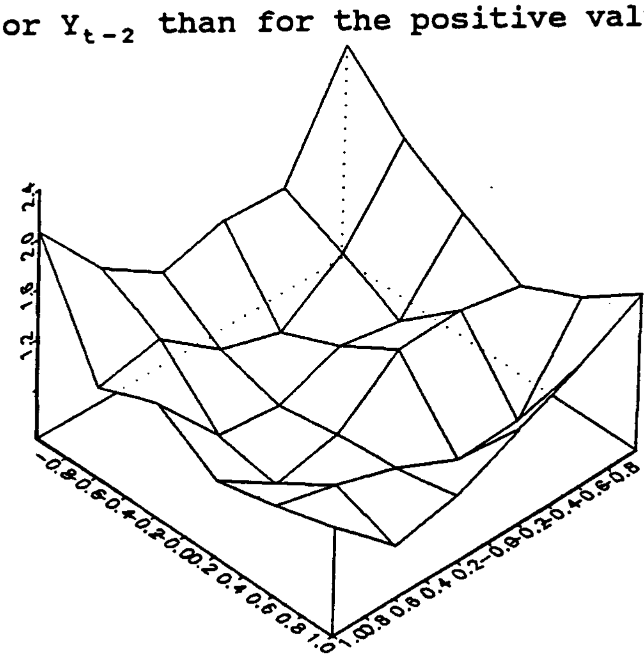

Figure 3

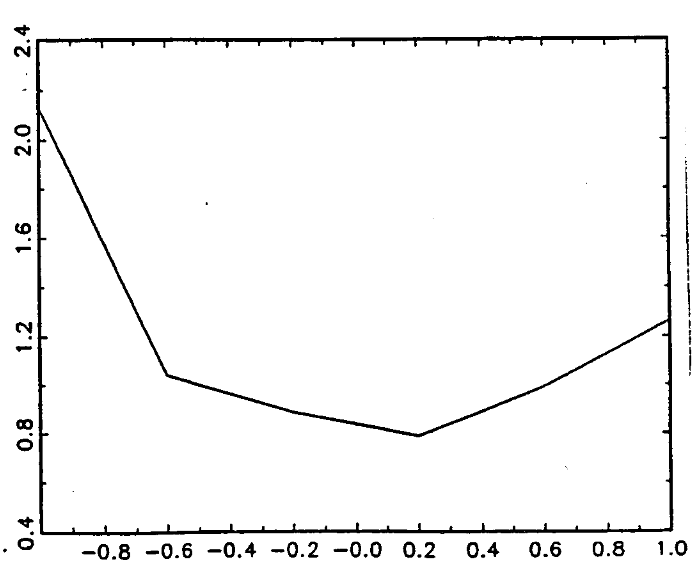

Figure 4

In order to study a possible influence of more than two lags and in order to keep a reasonable number of parameters, we now consider the following additive model for the conditional variance σ_t^2 :

(60) $$\sigma_t^2 = b_o + \sum_{i=1}^{3} \sum_{j=1}^{4} b_{ij} \mathbb{1}_{A_i} (Y_{t-j})$$

where the A_i , i=1,...,4 are the intervals defined by the boundary points -0,5%, 0, 0,5%. Note that the identifiability of the model is reached by imposing that the differential impact the fourth interval A_4 = [0,5%, +∞], is zero at all lags. The P.M.L estimates of the parameters are given in table 1, as well as the t-ratios based on the M.L. formulae (i.e. using the hessian of the Log-likelihood for computing the variances) and on the PML formulae.

	Estimates x 10^{-5}	t ratios (PML)	t ratios (ML)
b_o	11.9	7.2	8.4
b_{11}	5.1	2.0	2.6
b_{21}	- 3.3	- 2.0	- 2.7
b_{31}	- 2.0	- 1.4	- 1.6
b_{12}	6.4	2.0	3.1
b_{22}	- 1.6	- 1.1	- 1.4
b_{32}	- 2.3	- 1.3	- 1.7
b_{13}	6.6	2.3	3.3
b_{23}	0.2	0.1	0.2
b_{33}	- 0.7	- 0.5	- 0.7
b_{14}	6.4	2.2	2.9
b_{24}	- 1.7	- 1.0	1.3
b_{34}	- 3.2	- 2.5	- 3.0

Table 1
(Additive model)

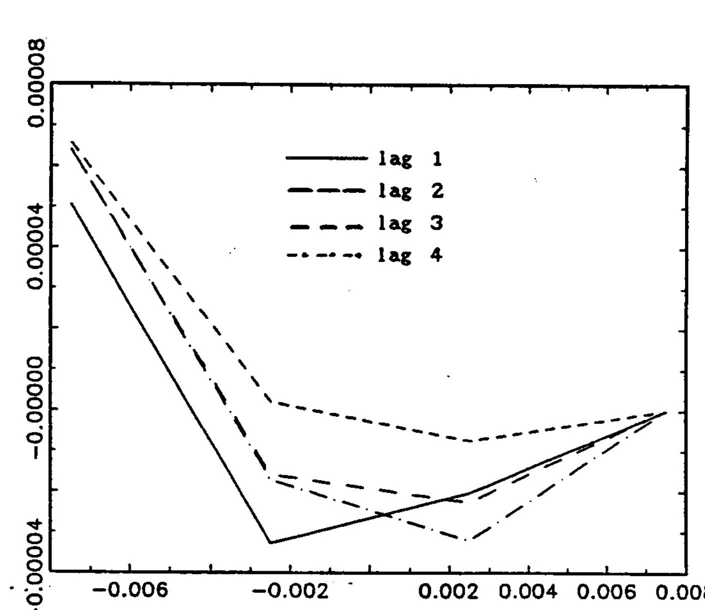

Figure 5
(Additive model)

As expected, the M.L. t-ratios are always too optimistic compared to the PML t-ratios. Note however that, according to both ML and PML t-ratios, the differential impacts of the first class A_1 = [-∞, 0.5%] with respect to the reference class A_4 = [0,5%, ∞], i.e. $\hat{b}_{11}$, $\hat{b}_{12}$, $\hat{b}_{13}$ $\hat{b}_{14}$, are significantly different from zero and positive for all lags. This is a strong confirmation of the non symmetrical effects of the past values of Y_t on the conditional variance.

Moreover figure 5 shows that the profiles of the reaction functions at different lags are similar. Following the previous remark it is natural to test, within the previous model, the restrictions implied by a G-QTARCH specification i.e. :

$$(61) \qquad \exists\lambda : b_{ij} = \lambda b_{i,j-1} \qquad \begin{array}{l} i = 1,2,3 \\ j = 2,3,4 \end{array}$$

A test, asymptotically equivalent to the (pseudo) Wald test, is easily implemented by using the A.L.S. theory. In a first step we get the OLS estimates $\tilde{\lambda}$ of λ in the linear model :

$$(62) \qquad \hat{b}_{ij} = \lambda\, \hat{b}_{i,j-1} + u_{ij} \qquad \begin{array}{l} i = 1,2,3 \\ j = 2,3,4 \end{array}$$

In a second step we apply the GLS method to the same linear model by using the covariance matrix of the u_{ij}'s:$V_{as}(\hat{b}_{ij}-\lambda\hat{b}_{i,j-1})_{\lambda=\tilde{\lambda}}$. The test statistics based on the PML and on the ML approaches are respectively :

$$\xi_{PML} = 2.8, \ \xi_{ML} = 5.8$$

Compared to the quantiles of the $\chi^2(8)$ distribution at any reasonable level, these statistics are not significant and the G-QTARCH specification is accepted.

Therefore we now estimate a G-QTARCH model ; since this kind of specification is parsimonious for the parameters describing the differences between the profiles at different lags, we can affect more parameters to a precise description of the within one lag feature of these profiles. More precisely we consider the model :

$$(63) \qquad \sigma_t^2 = \delta_0 \sigma_{t-1}^2 + \sum_{i=1}^{12} \delta_i \mathbb{1}_{A_i}(Y_{t-1})$$

where the A_i, i=1,2,...,12 are the intervals defined by the boundary points : -1.7%, -1.1%, -0.7%, -0.4%, -0.2%, 0%, 0.2%, 0.4%, 0.7%, 1.1%, 1.7%.

The results of the PML estimation are given in table 2 and the profiles of the δ_i, i=1,...,12 are shown in figure 6.

	Estimates $\times 10^{-5}$ except δ_o	t ratios (PML)	t ratios (ML)
δ_o	0.836	16.8	25.8
δ_1	14.06	2.7	4.5
δ_2	4.80	2.0	2.9
δ_3	2.69	1.7	2.3
δ_4	2.14	1.7	2.4
δ_5	0.16	0.2	0.3
δ_6	0.02	0.03	0.03
δ_7	0.03	0.06	0.07
δ_8	0.21	0.3	0.4
δ_9	1.51	1.8	2.4
δ_{10}	1.51	1.5	2.1
δ_{11}	1.31	1.2	1.4
δ_{12}	4.38	2.2	2.8

Table 2
(Generalized QTARCH model)

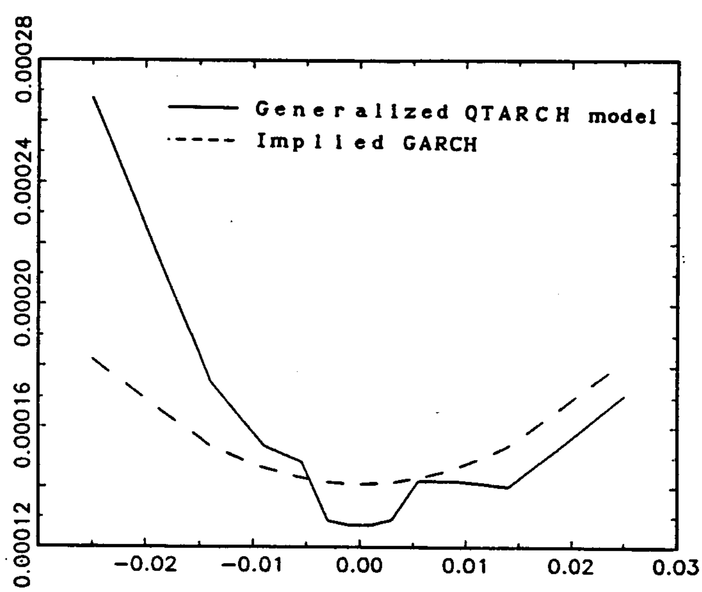

Figure 6
Graph of σ_t^2 as a function of Y_{t-1} (σ_{t-1}^2 equal to the marginal variance)

From table 2 it is seen that the autoregressive coefficient δ_0 is highly significant and that the δ_i corresponding to large values of $|Y_{t-1}|$ are also significantly different from zero ; on the contrary the δ_i corresponding to small valueS of $|Y_{t-1}|$ are not significantly different from zero. Moreover, the non symmetrical feature already mentioned is still particularly clear (see also figure 6).

It is now possible to test the restrictions implied by a GARCH formulation, i.e. :

(64) $\exists\ \lambda,\mu : \delta_i = \lambda + \mu\, a_i^2 \qquad i = 1,\ldots,12$

where a_i is the center of the interval A_i .

The A.L.S. approach provides the Wald tests based the PML and ML methods :

$$\xi_{PML} = 24.8 \qquad\qquad \xi_{ML} = 41.5$$

If we compare ξ_{PML} and ξ_{ML} to the quantiles of the $\chi^2(10)$ distribution, the GARCH specification is rejected at all reasonable levels, since $\chi^2_{0,95}(10) = 18.3$ and $\chi^2_{0,99}(10) = 23.2$. More precisely, figure 6 shows that the GARCH formulation could imply a serious distorsion of the parameter shape.

In order to smooth the shape of the response function to Y_{t-1} we have used the following non parametric technique. We first simulate σ_t^2 using model (63) (with σ_0^2 equal to the marginal variance 15.3 10^{-5}) then we apply a kernel regression technique (with a gaussian kernel) of $\hat{\sigma}_t^2$ on $\hat{\sigma}_{t-1}^2$ and y_{t-1}. The shape of the curve obtained is given in figures 7 and 8 (σ_{t-1}^2 is equal to the marginal variance 15.3 10^{-5}). On figure 7 y_{t-1} varies between -2.5% and 2.5% ; on figure 8, we have extended the range of y_{t-1} between -4% and 4%, however the shape for large values of $|y_{t-1}|$ becomes less precise because of the small number of observations (for instance 11 smaller than -3.5% and 5 greater than 3.5%). On these figures is also shown (dotted line) the parabola associated with the GARCH(1,1) model :

(65) $$\sigma_t^2 = 0.792\ \sigma_{t-1}^2 + 0.159\ y_{t-1}^2 + 7.7.10^{-6}$$

PML t ratios :	(13.2)	(3.0)	(2.3)
ML t ratios :	(19.8)	(5.0)	(3.2)

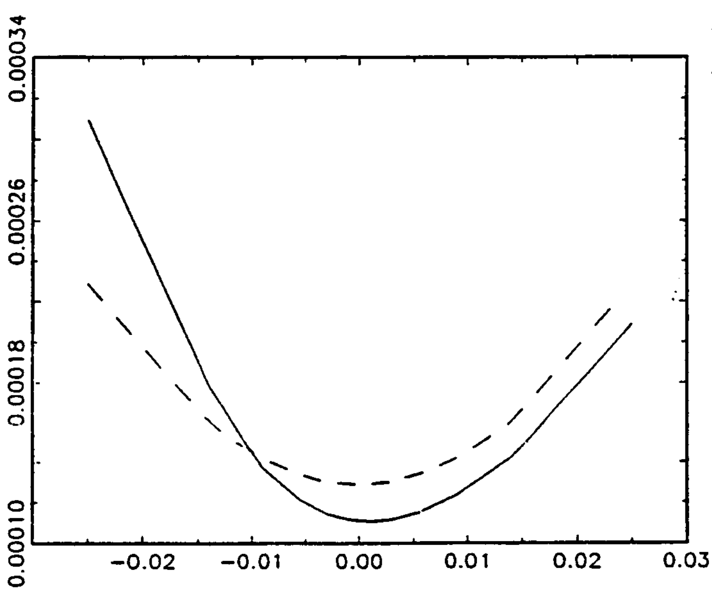

Figure 7

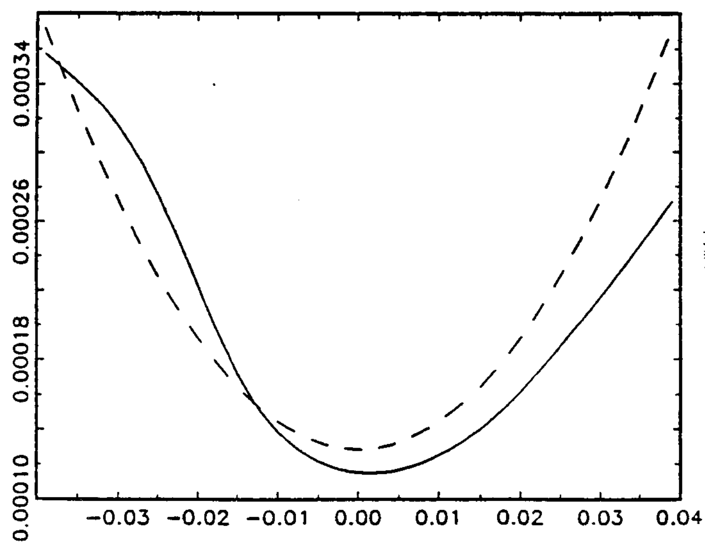

Figure 8

The curves of figures 7 and 8 again illustrate the non symmetry issue.

Such an analysis could obviously be pursued in various directions: specifications of parametric functional forms $\sigma_t^2=f(\sigma_{t-1}^2, y_{t-1})$, test of stability for different subperiods, specific effects of some days in the week, impact of outliers... For all these problems the statistical methods proposed in this paper are likely to be useful.

VI - CONCLUDING REMARKS

In this paper we have studied a class of conditionnally heteroscedastic models, called the QTARCH models, both in their probabilistic and their statistical aspects. This kind of models is easily implemented and seems to provide a flexible tool for a deep investigation of the conditional means and variances. These models can be used in a purely descriptive way or they can be used as a framework for testing successive restrictions based on statistical or economic considerations. Moreover, as shown in the application, our approach could be combined with non parametric techniques and it could be also useful for suggesting relevant parametric models.

Appendix 1
Spectral Decomposition

Let us consider the case of a stochastic matrix P which admits a diagonal representation. Since P is completely regular, 1 is a single eigenvalue and the other eigenvalues λ_j $j=1,...,J-1$ have a modulus strictly smaller than one. Therefore we have :

$$(*) \qquad P = M \begin{pmatrix} 1 & \lambda_1 & 0 \\ 0 & & \lambda_{J-1} \end{pmatrix} M^{-1} ,$$

where M is a complex matrix whose columns are eigenvectors of P. The first columns may be chosen as π and the other ones are denoted by $a_1,...a_{J-1}$. In a similar way $(M^{-1})'$ is a matrix whose columns are eigenvectors of P'. The first column may be chosen as e (since $\pi'e=1$) and the other ones are denoted by $b_1,...,b_{J-1}$.

With these choices and notations, equation (*) becomes :

$$P = \pi\, e' + \sum_{j=1}^{J-1} \lambda_j a_j b_j' ,$$

and the relationship $MM^{-1} = I$ gives the conditions

$$b_k' a_j = \delta_{kj} \quad k,j = 1,\ldots,J \text{ (with } a_j=\pi,\ b_j=e).$$

Appendix 2
Asymptotic Properties of the P.M.L. estimators

The asymptotic normality of the P.M.L. estimators and the asymptotic independence of the $\hat{\alpha}_j$'s, $\hat{\beta}_j$'s are classical results and we shall only focus on the derivation of the asymptotic covariance matrices.

i) Asymptotic covariance matrix of $\sqrt{T}(\hat{\alpha}_j-\alpha_j)$

From the stationarity properties, we deduce :

$$V_{as}\left[\sqrt{T}\begin{pmatrix}\frac{1}{T}\sum_t Y_t \mathbb{1}_{A_j}(Y_{t-1}) & - E\, Y_t \mathbb{1}_{A_j}(Y_{t-1}) \\ \frac{1}{T}\sum_t \mathbb{1}_{A_j}(Y_{t-1}) & - E\, \mathbb{1}_{A_j}(Y_{t-1})\end{pmatrix}\right]$$

$$= V_{as}\left[\sqrt{T}\begin{pmatrix}\frac{1}{T}\sum_t Y_t \mathbb{1}_{A_j}(Y_{t-1}) & - \alpha_j\pi_j \\ \frac{1}{T}\sum_t \mathbb{1}_{A_j}(Y_{t-1}) & - \pi_j\end{pmatrix}\right]$$

$$= \begin{bmatrix} V(Y_t \mathbb{1}_{A_j}(Y_{t-1})) & Cov(Y_t\mathbb{1}_{A_j}(Y_{t-1}),\ \mathbb{1}_{A_j}(Y_{t-1})) \\ Cov(\mathbb{1}_{A_j}(Y_{t-1}),\ Y_t\mathbb{1}_{A_j}(Y_{t-1})) & V(\mathbb{1}_{A_j}(Y_{t-1})) \end{bmatrix}$$

$$= \begin{bmatrix} \alpha_j\alpha_j'\pi_j(1-\pi_j) + \beta_j^2\pi_j & \alpha_j\pi_j(1-\pi_j) \\ \alpha_j'\pi_j(1-\pi_j) & \pi_j(1-\pi_j) \end{bmatrix}$$

Then we deduce the asymptotic covariance matrix of $\sqrt{T}(\hat{\alpha}_j-\alpha_j)$

where : $$\hat{\alpha}_j = \frac{\sum_t Y_t \mathbb{1}_{A_j}(Y_{t-1})}{\sum_t \mathbb{1}_{A_j}(Y_{t-1})}.$$

We get :

$$V_{as}[\sqrt{T}(\hat{\alpha}_j-\alpha_j)]$$

$$= \frac{1}{\pi_j^2}\left(\alpha_j\alpha'_j\pi_j(1-\pi_j) + \beta_j^2\pi_j\right) + \frac{\alpha_j\alpha'_j}{\pi_j^2}\pi_j(1-\pi_j) - \frac{2\,\alpha_j\alpha'_j\pi_j(1-\pi_j)}{\pi_j^2}$$

$$= \frac{\beta_j^2}{\pi_j}.$$

ii) <u>Asymptotic covariance matrix of</u> $\sqrt{T}[\mathrm{vec}\,\hat{\beta}_j^2 - \mathrm{vec}\,\beta_j^2]$

The asymptotic covariance matrix of $\sqrt{T}(\mathrm{vec}\,\hat{\beta}_j^2 - \mathrm{vec}\,\beta_j^2)$ can be derived assuming $\alpha_j=0$, $\forall j$ without loss of generality. In such a case the estimator is :

$$\hat{\beta}_j^2 = \frac{1}{T_j}\sum_{t\in B_j} Y_t Y'_t = \frac{\sum_t Y_t Y'_t \, \mathbb{1}_{A_j}(Y_{t-1})}{\sum \mathbb{1}_{A_j}(Y_{t-1})}$$

$$= \frac{\sum_t v_t v'_t \, \mathbb{1}_{A_j}(v_{t-1})}{\sum_t \mathbb{1}_{A_j}(v_{t-1})}$$

where : $$Y_t = v_t = \sum_{j=1}^{J} \beta_j \, \mathbb{1}_{A_j}(v_{t-1})\, u_t .$$

Let us denote by $\hat{B}_{k\ell j}$ the (k,l) entry of $\hat{\beta}_j^2$, we have :

$$B_{k\ell j} = \frac{\sum_t v_{kt} v_{\ell t} \, \mathbb{1}_{A_j}(v_{t-1})}{\sum_t \mathbb{1}_{A_j}(v_{t-1})}$$

The asymptotic covariance matrix of the $\hat{B}_{k\ell j}$'s, $k\leq l$ may be derived from the properties of the variables :

$$A_{k\ell j} = v_{kt} v_{\ell t} \, \mathbb{1}_{A_j}(v_{t-1}) \quad \text{and} \quad \mathbb{1}_{A_j}(v_{t-1})$$

a) We have :

$$E[v_t v'_t \, \mathbb{1}_{A_j}(v_{t-1})] = E[\beta_j u_t u'_t \beta_j \, \mathbb{1}_{A_j}(v_{t-1})] = \beta_j^2 \pi_j .$$

We deduce that $E(A_{k\ell j}) = B_{k\ell j}\pi_j$.

b) Let us now compute the variances and covariances. We have :

$$\mathrm{Cov}(A_{klj}, A_{k^*l^*j})$$

$$= \mathrm{Cov}[E(A_{klj}/v_{t-1}), E(A_{k^*l^*j}/v_{t-1})]$$

$$+ E[\mathrm{Cov}(A_{klj}, A_{k^*l^*j}/v_{t-1})]$$

$$= \mathrm{Cov}[E(\beta_j^{k\prime} u_t u'_t \beta_j^{\ell}) \, \mathbb{1}_{A_j}(v_{t-1}), E(\beta_j^{k^*\prime} u_t u'_t \beta_j^{l^*}) \, \mathbb{1}_{A_j}(v_{t-1})]$$

$$+ E[\mathbb{1}_{A_j}(v_{t-1}) \,\mathrm{Cov}(\beta_j^{k\prime} u_t u'_t \beta_j^{\ell}, \beta_j^{k^*\prime} u_t u'_t \beta_j^{l^*})],$$

where β_j^k is the k^{th} column of β_j. Noting that β_j is a symmetric matrix, we see that $\beta_j^{k\prime}\beta_j^{\ell}$ is the (k,l) element of β_j^2, i.e. $B_{k\ell j}$. Therefore we get :

$$\text{Cov}(A_{klj}, A_{k^*l^*j}) = B_{klj}B_{k^*l^*j'}\ \pi_j(1-\pi_j)$$

$$+ \pi_j\ \text{Cov}(\beta_j^{k\prime}u_t u_t'\beta_j^{\ell},\ \beta_j^{k^*\prime} u_t u_t'\beta_j^{l^*})$$

$$= B_{klj}B_{k^*l^*j}\pi_j(1-\pi_j)+\pi_j C_{klk^*l^*}\ \text{(say)}$$

Similarly, we get :

$$\text{Cov}(A_{klj}, \mathbb{1}_{A_j}(v_{t-1}))$$

$$= \text{Cov}(B_{klj}\mathbb{1}_{A_j}(v_{t-1}), \mathbb{1}_{A_j}(v_{t-1})) + 0$$

$$= B_{klj}\ \pi_j(1-\pi_j)\ .$$

c) Now, we may apply the δ method to derive the asymptotic covariance :

$$\text{Cov}_{as}[\sqrt{T}(\hat{B}_{klj}-B_{klj}), \sqrt{T}(\hat{B}_{k^*l^*j}-B_{k^*l^*j})]$$

$$= \frac{1}{\pi_j^2}\ \text{Cov}(A_{klj} - \mathbb{1}_{A_j}(v_{t-1})\ B_{klj}, A_{k^*l^*j} - \mathbb{1}_{A_j}(v_{t-1})\ B_{k^*l^*j})$$

$$= \frac{1}{\pi_j^2}\ \{B_{klj}B_{k^*l^*j}\pi_j(1-\pi_j) + \pi_j C_{klk^*l^*}$$

$$- 2B_{klj}B_{k^*l^*j}\pi_j(1-\pi_\ell) + B_{klj}B_{k^*l^*j}\pi_j(1-\pi_j)\}$$

$$= \frac{1}{\pi_j}\ C_{klk^*l^*}$$

iii) <u>Expression of</u> $C_{klk^*l^*}$ <u>in the univariate case</u>

We get :

$$C_{klk^*l^*} = C = V(\beta_j^2 u_t^2)$$

$$= \beta_j^4[Eu_t^4-(Eu_t^2)^2]$$

$$= \beta_j^4(\mu_4-1)\ .$$

REFERENCES

Bollerslev, T. (1986) : "Generalized Autoregressive Conditional Heteroskedasticity", Journal of Econometrics, 37, 307-327.

Chan, K.S., Petrucelli, Tong, H. and **Woolford (1985)** : "A Multiple Threshold AR(1) Model", Journal of Applied Probability, 2, 267-279.

Chan, K.S. and **Tong, H. (1986)** : "On Estimating Thresholds in Autoregressive Models", Journal of Time Series Analysis, 7, 179-190.

Cox, J., Ross, S. and **Rubinstein, M. (1979)** : "Option Pricing : A Simplified Approach", Journal of Financial Economics, 7, 229-263.

Cox, J. and **Rubinstein, M. (1985)** : "Option Markets", Prentice Hall, New Jersey.

Diebold, F. and **Nerlove, M. (1986)** : "Factor Structure in a Multivariate GARCH Model of Exchange Rate Fluctuations", Univ. of Pennsylvania DP.

Diebold, F. and **Nerlove, M. (1989)** : "The Dynamics of Exchange Rate Volatility : A Multivariate Latent Factor ARCH Model", Journal of Applied Econometrics 4, 1-22.

Engle, R.F. (1982) : "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of U.K. Inflation", Econometrica, 50, 987-1008.

Engle, R.F. (1987) : "Multivariate ARCH with Factor Structure : Cointegration in Variance", U.C.S.D., D.P.

Engle, R.F. and **Bollerslev, T. (1986)** : "Modeling the Persistence of Conditional Variances", Econometric Reviews, 5, 1-50.

Engle, R.F. and **Gonzales-Rivera, G. (1989)** : "Semiparametric ARCH Models", U.C.S.D. Discussion Paper.

Engle, R., Ng, V. and **Rothschild, M. (1989)** : "Asset Pricing with a Factor ARCH Covariance Structure : Empirical Estimates for Treasury Bills", U.C.S.D., D.P.

Engle, R.F., Lilien, D. and **Robbins, R. (1987)** : "Estimating Time Varying Risk Premia in the Term Structure : the ARCH-M Model", Econometrica, 55, 391-407.

Gallant, R. and **Tauchen, G. (1989)** : "Seminonparametric Estimation of Conditionally Constrained Heterogeneous Processes : Asset Pricing Applications", Econometrica, 57, 1091-1120.

Gantmacher, F.R. (1966) : Théorie des Matrices, Dunod, Paris.

Gourieroux, C., Monfort, A. and Trognon, A. (1984) : "Pseudo Maximum Likelihood Methods : Theory", Econometrica, 52, 681-700.

Gourieroux, C., Monfort, A. and Trognon, A. (1985) : "Moindres carrés asymptotiques", Annales de l'INSEE, n°58, 91-122.

Gourieroux, C. and **Monfort, A. (1989-a)** : Statistique et modèles économétriques, 2 volumes, Economica, Paris.

Gourieroux, C. and **Monfort, A. (1989-b)** : "A General Framework for testing a Null Hypothesis in a Mixed Form", Econometric Theory, 5, 63-82.

Gourieroux, C., Monfort, A. and Renault, E. (1988) : "Contraintes bilinéaires : Estimation et test", in Mélanges Economiques, Essais en l'honneur d'E. MALINVAUD, Economica, Paris.

Gregory, A.W. (1989) : "A Nonparametric Test for Autoregressive Condiditional Heteroscedasticity : A Markov Chain Approach", Journal of Business and Economic Statistics, 7, 107-115.

Hendry, D. and Mizon, G. (1990) : "Evaluating Dynamic Econometric Models by Encompassing the VAR", Oxford discussion paper.

Huang, C.F. and **Litzenberger, R.M. (1988)** : Foundations for Financial Economics, North-Holland.

Melard, G. and **Roy, R. (1987)** : "Modèles de séries chronologiques avec seuils", to appear in Statistique et Analyse des Données.

Monfort, A. and Rabemananjara, R. (1990) : "From a VAR model to a structural model, with an application to the wage-price spiral", Journal of Applied Econometrics, 5, 203-227.

Nerlove, M., Diebold, F., Van Beek, H. and **Cheung, Y. (1988)** : "A Multivariate ARCH Model of Foreign Exchange Rate Determination", Univ. of Pennsylvania D.P.

Saikkonen, P. and **Luukkonen (1986)** : "Lagrange Multiplier Tests for Testing Nonlinearities in Time Series Models", Dept of Statistics, Univ. of Helsinki.

Szroeter, J. (1983) : "Generalized Wald Methods for Testing Nonlinear Implicit and Overidentifying Restrictions", Econometrica, 51, 335-348.

Tong, H. (1983) : "Threshold Models in Nonlinear Time Series", Edited by Brillinger, D. Fienberg, S., Gani, J. Hartigan, J., Krickeberg, K., Springer Verlag.

Tong, M. and **Lim, K.S. (1980)** : "Threshold Autoregressions, Limit Cycles and Cyclical Data", J.R.S.S., B. 42, 245-292.

White, H. (1980) : "A Heteroskedastic Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity", Econometrica, 48, 817-838.

Zakoian, J.M. (1990) : "Threshold ARCH Models", document de travail ENSAE/INSEE.