Test de changement de régimes dans des séries financières par un modèle conditionnellement hétéroscédastique à seuil endogène

Youssef Saïdi

Département de la Recherche, Bank Al-Maghrib, Rabat, Maroc,
E-mails: saidiyoussef@hotmail.com / y.saidi@bkam.ma

Résumé. En finance, les modèles autorégressifs conditionnellement hétéroscédastiques (ARCH), et leurs nombreuses extensions se sont avérés être des instruments très efficaces. Une nouvelle classe de modèles conditionnellement hétéroscédastiques non linéaires, introduite dans Saïdi (2003) et Saïdi et Zakoïan (2006), fait dépendre la volatilité de la position relative des innovations passées. Cette dernière se rattache aux extensions précédentes par l'existence de plusieurs régimes. Dans ce papier, nous proposons une méthode de test de changement de régime dont la construction repose sur la nouvelle classe de modèles introduite par Saïdi (2003) et Saïdi et Zakoïan (2006), et inspirée de la méthode développée par Tsay (1989) pour les modèles autorégressifs à seuils. Ensuite, nous testons la présence de la modification des régimes de volatilité dans le rendement de l'indice boursier CAC 40 en utilisant le modèle proposé.

Mots-clés. Modèles ARCH, modèles à seuils, séries temporelles non linéaires, test de changement de régimes, maximum de vraisemblance.

Abstract : In finance, the autoregressive conditional heteroscedasticity (ARCH) models and their many extensions proved to be very effective instruments. In Saïdi (2003) and Saïdi & Zakoïan (2006), a new class of conditional heteroskedastic nonlinear models was developed and studied. The volatility of the variable, at time t, depends on the relative position of the past variables. Our approach in the extension of this model is related to the previous studies devoted to the existence of several regimes in time series.
In this paper, we propose a test statistic based on the methods presented in Tsay (1989) for the threshold autoregressive model. Thereafter, we test the existence of two possible volatility regimes (high volatility and low) of the returns of the CAC 40 index by using the proposed model.

Keywords : ARCH models, threshold models, nonlinear time series, test of changes in regime, maximum likelihood.

1 Introduction

Depuis l'article d'Engle (1982) introduisant les modèles ARCH (autorégressifs conditionnellement hétéroscédastiques), de nombreuses spécifications de la volatilité ont été proposées. Le modèle linéaire GARCH(p,q) généralisé par Bollerslev (1986) se présente sous la forme :

$$\left\{ \begin{array}{ll} \varepsilon_t & = h_t^{1/2}\eta_t \ , \\ h_t & = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j}, \end{array} \right. \tag{1}$$

où (η_t) est une suite de variables indépendantes, identiquement distribuées (i.i.d.) centrées et de variance unité, les α_i et β_j sont des constantes positives et ω est une constante strictement positive.

De nombreuses études empiriques ont démontré diverses limitations des modèles GARCH linéaires. Une première classe d'extensions a visé la prise en compte d'asymétries (voir Nelson (1991), modèle GARCH exponentiel; Zakoïan (1994), modèle GARCH à seuil, TGARCH) dans les modèles GARCH. En finance, l'accroissement de volatilité dû à une baisse des prix est généralement supérieur à celui résultant d'une hausse de même ampleur. Une formulation, faisant dépendre la volatilité du signe des valeurs passées, a été introduite dans le modèle TGARCH (p,q) :

$$h_t = \omega + \sum_{i=1}^q \alpha_{i,+}\varepsilon_{t-i}^+ - \alpha_{i,-}\varepsilon_{t-i}^- + \sum_{j=1}^p \beta_j h_{t-j},$$

où $\varepsilon_t^+ = \max(\varepsilon_t, 0)$, $\varepsilon_t^- = \min(\varepsilon_t, 0)$, les constantes $\alpha_{i,+}, \alpha_{i,-}, \beta_j$ sont positives et ω strictement positive.

Une autre classe de généralisation permettant de dissocier les dynamiques des grandes et petites innovations est celle des modèles GARCH à changements de régimes markoviens (voir Hamilton et Susmel (1994), Francq, Roussignol et Zakoïan (2001)).

Le modèle, que nous proposons, se rattache aux extensions précédentes par l'existence de plusieurs régimes. Il a pour but de prendre en compte des dynamiques plus complexes, par exemple, l'impact de ε_{t-1}^2 sur la volatilité de la date t dépendant de la position relative de certaines variables antérieures à la date t. Ce n'est donc pas le niveau des variations de prix qui gouverne les changements de régime mais le fait qu'une variation soit grande, ou petite, par rapport aux variations des dates précédentes.

2 Modèle et propriétés

On considère dans ce papier le modèle d'ordre 1 défini par :

$$\left\{ \begin{array}{l} \varepsilon_t :=: h_t^{1/2}\eta_t \ , \\ h_t = \omega + \left(\sum_{i=1}^r \alpha_i \mathbb{1}_{Z_{t-1}\in R_i}\right) \varepsilon_{t-1}^2, \end{array} \right. \tag{2}$$

où $(\eta_t)_{t\in \mathbb{Z}}$ est une suite de variables i.i.d centrées et de variance 1, définies sur un espace probabilisé $(\Omega \ ,\mathcal{A}, \ \mathbb{P} \ )$; Z_{t-1} est une variable à valeurs dans $\mathbb{R}^d$ $(d \in \{1,2,\ldots\})$ appartenant à la tribu $\mathcal{F}_{t-1}$ engendrée par $\{\varepsilon_{t-j}, j>0\}$; $(R_i)_{i=1,\ldots,r}$ est une partition de $\mathbb{R}^d$; $\omega > 0$ et pour $i = 1,\ldots,r$, $\alpha_i \geq 0$. On suppose que η_t est indépendante de $\{\varepsilon_{t-j}, j>0\}$ et on note $\mu_{2m} = E(\eta_t^{2m})$. Dans la suite de ce papier, nous allons appeler le modèle défini par (2) *un ARCH à seuils endogènes* et sera noté ETARCH$(1,r)$ (Endogenous Threshold ARCH en anglais).

Le modèle ARCH(1) standard est obtenu lorsque les α_i sont égaux ou lorsque $r = 1$. Le modèle TARCH(1) introduit par Rabemananjara et Zakoïan (1993), correspond à $Z_{t-1} = \varepsilon_{t-1}$, $r = 2$, $R_1 = \mathbb{R}^+$ et $R_2 = \mathbb{R}^-$. Un cas particulier *à deux régimes* de notre modèle est obtenu en choisissant $Z_{t-1} = \varepsilon_{t-1}^2/\varepsilon_{t-2}^2$, $r = 2$, $R_1 = [0, k]$, $R_2 =]k, +\infty[$ avec $k \geq 0$:

$$\begin{cases} \varepsilon_t :=: h_t^{1/2}\eta_t \,, \\ h_t = \omega + \alpha_1\varepsilon_{t-1}^2 \mathbb{1}_{[\varepsilon_{t-1}^2 \leq k\varepsilon_{t-2}^2]} + \alpha_2\varepsilon_{t-1}^2 \mathbb{1}_{[\varepsilon_{t-1}^2 > k\varepsilon_{t-2}^2]}. \end{cases} \tag{3}$$

Dans le modèle ETARCH$(1, 2)$ défini par (3), deux régimes conditionnellement hétéroscédastiques correspondant respectivement aux 'petites' et aux 'grandes' variations sont représentés dès que $\alpha_1 \neq \alpha_2$ et $k > 0$. Les régimes 1 et 2 supposent que la volatilité est modérée au cours du temps (périodes de calme) mais lorsque la volatilité a augmenté consécutivement dans le passé récent, alors on peut s'attendre à ce que la volatilité augmente encore plus dans le présent (périodes d'agitation). En quelque sorte, le modèle exprime que les marchés sont "excités" si un certain seuil a été dépassé par les accroissements de rentabilités passées.

Dans Saïdi (2003), une étude probabiliste et statistique a été réalisée pour la famille des modèles ETARCH(1,r) [1]. Ensuite, Saïdi et Zakoïan (2006) améliorent et renforcent les résultats de l'étude probabiliste élaborée par Saïdi (2003).

3 Test d'existence de plusieurs régimes

L'objectif de cette partie est de présenter une méthode de test d'existence de plusieurs régimes selon le modèle ETARCH$(1, 2)$ défini par (3). Ensuite, une étude par simulation de la puissance et le niveau du test est réalisée.

3.1 Méthode du test

Nous proposons d'adapter la méthode de Tsay (1989) développée initialement pour des modèles à seuil (TAR) sur la moyenne. Cette méthode ne nécessite pas d'estimer le modèle contraint (plusieurs régimes) et donc de connaître les seuils. Le but est de construire un test d'hypothèse de linéarité (modèle ARCH(1) classique) contre celle de présence de seuils (modèle de type (2)). Pour simplifier, nous supposons que $r = 2$, $R_1 = [0, k]$, $R_2 =]k, +\infty[$ avec $k \geq 0$ et nous réécrivons (2) de la façon suivante :

$$\varepsilon_t = \sqrt{h_t}\eta_t, \qquad\qquad h_t = \begin{cases} \omega_0 + \alpha_{01}\varepsilon_{t-1}^2 & \text{si} \quad Z_{t-1} \leq k, \\ \omega_0 + \alpha_{02}\varepsilon_{t-1}^2 & \text{si} \quad Z_{t-1} > k. \end{cases} \tag{4}$$

Nous rappelons que (Z_t) est une variable à valeurs dans $\mathbb{R}^+$ appartenant à la tribu $\mathcal{F}_{t-1}$ engendrée par $\{\varepsilon_{t-j}, j > 0\}$. Par exemple, $Z_t = \frac{\varepsilon_t^2}{\varepsilon_{t-1}^2}$, ou $Z_t = \varepsilon_t^2$.

[1]Il s'agit des conditions de stationnarité et d'existence de moments du modèle ETARCH$(1, 2)$ avec $\alpha_1 = 0$. Ainsi, trois méthodes d'estimation ont été comparées sur le modèle plus général ETARCH$(1, r)$ à savoir la méthode des MCO, la méthode des moments et la méthode du maximum de vraisemblance.

Supposons que l'on dispose des observations $\varepsilon_1, \cdots, \varepsilon_n$; ainsi que de n_0 valeurs initiales, qui permettent d'obtenir les Z_t, $t = 1, \cdots, n$. Le nombre n_0 dépendra de la fonction Z_t. Notre approche consiste à tester

$$H_0 : \alpha_{01} = \alpha_{02} \quad \text{ou} \quad k = 0 \quad \text{contre} \quad H_a : \alpha_{01} \neq \alpha_{02} \quad \text{et} \quad k > 0.$$
$$\text{(un seul régime)} \qquad\qquad \text{(deux régimes)}$$

Pour construire le test, nous considérons la statistique d'ordre sur (Z_t) et notons $Z_{(i)}$ le i^{eme} plus petit élément. Posons $\tau(i)$ la date (aléatoire) correspondant à $Z_{(i)}$:

$$\forall i = 1, 2, ..., n\,, \qquad \exists!\ t = \tau(i), \qquad Z_t = Z_{(i)}.$$

La statistique d'ordre sur Z_t permet de ranger les ε_t dans l'ordre suivant : $\varepsilon_{\tau(1)}, \cdots, \varepsilon_{\tau(n)}$. Avec ces notations, les équations du modèle (4) s'écrivent, pour ces observations, sous la forme :

$$\forall\ i = 1, 2, ..., n, \qquad \begin{cases} \varepsilon_{\tau(i)} \ = \ h^{\frac{1}{2}}_{\tau(i)} \eta_{\tau(i)}\ , \\ h_{\tau(i)} = \omega_0 + \alpha_{01}\varepsilon^2_{\tau(i)-1} \quad \text{si} \ \ i \leq s, \\ h_{\tau(i)} = \omega_0 + \alpha_{02}\varepsilon^2_{\tau(i)-1} \quad \text{si} \ \ i > s, \end{cases} \tag{5}$$

où s satisfait $Z_{(s)} \leq k < Z_{(s+1)}$. C'est un classement tel que les s premières observations $(\varepsilon_{\tau(i)})$ sont dans le premier régime et les $n-s$ suivantes dans le second. Pour ce classement, seul le nombre s dépend du seuil k. Considérons le modèle de régression :

$$\varepsilon^2_{\tau(i)} = \omega + \alpha\varepsilon^2_{\tau(i)-1} + \nu_{\tau(i)}, \qquad\qquad i = 1, ..., n. \tag{6}$$

Soit $\hat{\theta}^{(m)} = (\hat{\omega}^{(m)}, \hat{\alpha}^{(m)})$ l'estimateur des moindres carrés de $\theta = (\omega, \alpha)$ obtenu à partir des m premières valeurs des $(\varepsilon^2_{\tau(i)})$. L'idée du test est que si (ε_t) est de type (4), quand $i \leq s$ les résidus $(\nu_{\tau(i)})$ constituent approximativement un bruit blanc donc sont (approximativement) orthogonaux aux valeurs passées $\varepsilon^2_{\tau(i)-1}$ ($\nu_{\tau(i)}$ non corrélé avec $\varepsilon^2_{\tau(i)-1}$) . Quand m excède s, les résidus $(\nu_{\tau(i)})$ vont être biaisés car le régime change à l'indice $(s+1)$. Donc l'orthogonalité entre les résidus $(\nu_{\tau(i)})$ et $(\varepsilon^2_{\tau(i)-1})$ va se détruire progressivement dès que le nombre d'observations m prend des valeurs supérieures à s.
Les étapes du test sont les suivantes :

Étape 1 : Reclassement des observations selon la statistique d'ordre sur Z_t et le calcul des estimateurs $\hat{\theta}^{(m)} = (\hat{\omega}^{(m)}, \hat{\alpha}^{(m)})$, $m = m_0$, $m_0 + 1$, ...,$n - 1$, pour une valeur m_0 arbitraire suffisamment grande, par la méthode des moindres carrés.

Étape 2 : Calcul des résidus $\hat{\nu}_{\tau(m+1)} = \varepsilon^2_{\tau(m+1)} - \hat{\omega}^{(m)} - \hat{\alpha}^{(m)}\varepsilon^2_{\tau(m+1)-1}$.

Étape 3 : Régression, par la méthode des moindres carrés,

$$\hat{\nu}_{\tau(l)} = \Psi_0 + \Psi_1\varepsilon^2_{\tau(l)-1} + w_{\tau(l)}, \qquad l = m_0 + 1, m_0 + 2, ..., n \tag{7}$$

et calcul des résidus $\hat{w}_{\tau(l)}$.

Étape 4 : Test de l'hypothèse nulle H_0': $\Psi_0 = \Psi_1 = 0$ contre l'hypothèse alternative H_a': $\Psi_0 \neq 0$ ou $\Psi_1 \neq 0$.

3.2 Simulations

Nous utilisons 100 simulations, de taille $n = 2400$, du modèle (4) avec $Z_t = \frac{\varepsilon_t^2}{\varepsilon_{t-1}^2}$ et $(k, \omega, \alpha_1, \alpha_2) = (4, 0.1, 0.1, 0.6)$ (pour une étude de puissance) et d'un ARCH(1) avec $(\omega, \alpha) = (0.1, 0.2)$ (pour une étude de niveau). Nous appliquons les étapes 1, 2, 3 du test précédent pour $m_0 = 500$. Cette valeur est suffisamment grande pour fournir des estimateurs assez précis dès les premières valeurs.
Ainsi, nous constatons que la fréquence empirique de rejet sous H_0 dépasse la valeur nominale (1% ou 5%). Sous l'alternative H_a, cette fréquence qui est un estimateur de la puissance est très élevée.

4 Résultats Empiriques: l'indice CAC 40

Nous utilisons une série journalière des rendements du CAC 40 (logarithme du rapport du prix à la date t sur le prix à la date $t-1$) qui correspond à 2385 observations.

4.1 Test de changements de régime

Nous nous plaçons dans le cadre du modèle (4) avec $Z_t = \frac{\varepsilon_t^2}{\varepsilon_{t-1}^2}$. Nous commençons par l'application du test de changement de régime introduit ci-dessus avec $m_0 = 500$. En utilisant la statistique

$$\hat{F} = \frac{\sum \hat{\nu}_t^2 - \sum \hat{w}_t^2}{\sum \hat{w}_t^2} \; \frac{n - m_0 - 2}{2},$$

nous trouverons $\hat{F} = 11.83$. Pour un niveau de risque $\alpha = 0.01$ on a : $F_{1-\alpha}(2, n-m_0-2) = F_{0.99}(2, 1781) = 4.61$. Donc, d'après la règle de décision, H_0 est largement rejetée au profit de H_a. Ce qui nous permet de conclure l'existence de changements de régime de type (4).

4.2 Estimation des paramètres

Afin de localiser k, nous représentons graphiquement les estimations de $\hat{\omega}^{(m)}$ et $\hat{\alpha}^{(m)}$ de l'étape 1 du test. Nous observons une nette décroissance des estimations de α à partir des valeurs de m proches de 1600, correspondant à $Z_{(i)} = 3.22$. Ceci nous permet de localiser k avant cette valeur.
Afin d'affiner la localisation, nous balayons la région $Z_{(i)} \in [1.23, 3.21]$ en calculant les estimations de ω, α_1 et α_2 par la méthode du maximum de vraisemblance. Nous choisissons, selon le premier balayage, $\hat{k} = 2.5$ qui correspond à la plus grande vraisemblance.
Nous affinons le balayage autour de cette valeur avec une précision de l'ordre 10^{-2} et nous constatons que cette valeur a effectivement la plus grande vraisemblance. Nous estimons, ensuite, par la méthode du maximum de vraisemblance, le modèle (4) pour $k = 2.5$. Dans le Tableau 1, nous donnons les estimations des paramètres ω, α_1 et α_2 ainsi que leurs caractéristiques.

Tableau 1: Rendements du CAC 40 - Estimation par MV des paramètres du modèle (4)

Paramètres	Estimation	Ecart-type estimé [RATS]	T-stat	Signif
$k = 2.5$				
ω	0.0001	$0.26 * 10^{-5}$	41.97	0.00
α_1	0.30	$0.49 * 10^{-1}$	6.12	0.00
α_2	0.07	$0.16 * 10^{-1}$	4.43	$9.6 * 10^{-6}$

Afin de vérifier la plausibilité du modèle (4) estimé, nous le comparons avec d'autres modèles ARCH en utilisant des simulations.
En utilisant la série des rendements du CAC 40, nous avons estimé ainsi par la méthode du maximum de vraisemblance deux spécifications ARCH(2) et GARCH(1,1) et nous avons comparé les séries des simulations des trois spécifications avec celle du CAC 40. Comme l'on pouvait s'y attendre, les autocorrélations empiriques des quatre séries sont proches de 0. En utilisant les autocorrélations des carrés d'ordre 1, les résultats de comparaison indiquent que le modèle (4) a un ajustement meilleur de la série réelle en comparant sa performance avec celle des autres modèles.

Bibliographie

[1] Bollerslev, T. (1986), *Generalized Autoregressive Conditional Heteroskedasticity,* Journal of Econometrics, 31, 307–327.
[2] Engle, R. F., (1982), *Autoregressive Conditional Heteroskedasticity With Estimates of the Variance of the United Kingdom Inflation,* Econometrica, 50, 987–1007.
[3] Francq, C., Roussignol, M. et J-M. Zakoïan (2001), *Conditional Heteroskedasticity Driven by Hidden Markov Chains,* Journal of Time Series Analysis 22 (2), 197-220.
[4] Francq, C. & Zakoïan, J-M. (2010), *GARCH Models: Structure, Statistical Inference and Financial Applications.* John Wiley.
[5] Hamilton, J. D. et R. Susmel (1994), *Autoregressive Conditional Heteroskedasticity and Changes in Regime*, Journal of Econometrics, 64, 307-333.
[6] Saïdi, Y. (2003), *Étude probabiliste et statistique de modèles conditionnellement hétéroscédastiques non linéaires,* Thèse de doctorat de l'Université Lille 3. Disponible sur http://documents.univ-lille3.fr/files/pub/www/recherche/theses/saidi-youssef.pdf.
[7] Saïdi, Y. & Zakoïan, J-M. (2006), *Stationarity and geometric ergodicity of a class of nonlinear ARCH models,* Ann. Appl. Probab. 16, no. 4, 2256–2271.
[8] Tsay R.S (1989),*Testing and Modeling Threshold Autoregressive Processes*, Journal of American Statistical Association, Vol. 84, No. 405, 231-240.
[9] Zakoïan, J-M. (1994). Threshold heteroskedastic models. J. Econom. Dynam. Control 18 931–955.